

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
11 April 2002 (11.04.2002)

PCT

(10) International Publication Number
WO 02/29059 A2(51) International Patent Classification⁷: C12N 15/12,
C07K 14/70506437 (US). CATE, Richard, L. [US/US]; 40B Nichols
Road, Cohasset, MA 02025 (US). SAH, Dinah, W., Y.
[US/US]; 4 Longfellow Place, Apt. 2608, Boston, MA
02114 (US).

(21) International Application Number: PCT/US01/31488

(22) International Filing Date: 6 October 2001 (06.10.2001)

(74) Agents: HALEY, James, F. et al.; Fish & Neave, 1251
Avenue of the Americas, New York, NY 10020 (US).

(25) Filing Language: English

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU,
ZA, ZW.

(26) Publication Language: English

(30) Priority Data: 60/238,361 6 October 2000 (06.10.2000) US

(71) Applicants (for all designated States except US): YALE
UNIVERSITY [US/US]; 541 College Street, New Haven,
CT 06520 (US). BIOGEN, INC. [US/US]; 14 Cambridge
Center, Cambridge, MA 02142 (US).(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): STRITTMATTER,
Stephen, M. [US/US]; 96 Tulip Tree Drive, Guilford, CT

[Continued on next page]

(54) Title: NOGO RECEPTOR HOMOLOGS

	1				50
Human NOGO-R1	-----	-----	-----	-----	-----
Murine NOGO-R1	-----	-----	-----	-----	-----
Murine NOGO-R3	-----	-----	-----	-----	-----
Human NOGO-R3	-----	-----	-----	-----	-----
Human NOGO-R2	-----	-----	-----	-----	-----
Consensus	-----	-----	-----	-----	-----
	51	LRR NT	LRR 1		100
Human NOGO-R1	PKVTTSCTPQ	GLQAVFVGIP	AASQRIFLHG	NRISHVPAAS	FRACRNLTL
Murine NOGO-R1	PKVTTSCTPQ	GLQAVPTGIP	ASSQRIFLHG	NRISHVPAAS	FQSCRNLTL
Murine NOGO-R3	P.MTVSCQAH	NFAAIPGEGIP	EDSERIFLQ	NRITPLQGH	FSP.AMVTL
Human NOGO-R3	-----	-----	-----	-----	-----
Human NOGO-R2	P.PTVSCQAN	NFSSVPLSLP	PSTQRLFLQ	NRITLTPGT	FGS.NLLTL
Consensus	P--T-SC--	P---P---	---R-FL--	N-I-----	P-----L
	101	LRR 2	LRR 3		150
Human NOGO-R1	WLHSNVLARI	DAAAFGLAL	LEQLDLSDNA	QLRSVDPATF	HGLRLTLH
Murine NOGO-R1	WLHSNALARI	DAAAFGLTL	LEQLDLSDNA	QLHVVDPTTF	HGLSHLTLH
Murine NOGO-R3	WIYSNNITFI	APNTFEGFVH	LEELDLDGMR	QLRTLAPETP	QGLVKLHALY
Human NOGO-R3	WIYSNNITFI	HPSTFEGFVH	LEELDLDGMR	QLRTLAPETP	QGLVKLHALY
Human NOGO-R2	WLFSNNLSTI	YPGTFRHLQ	LEELDLDGMR	HLRSLEPTTF	QGLERLQSLH
Consensus	W--SN-----	I-----P-----	LE-LDL-DN-	-L-----P-TF	-GL--L--L-
	151	LRR 4	LRR 5		200
Human NOGO-R1	LDRCGLELQ	PGLFRGLAAL	QYLYLDQNAL	QALPDDTFRD	LGNLTHLPLH
Murine NOGO-R1	LDRCGLELQ	PGLFRGLAAL	QYLYLDQNNL	QALPDNTFRD	LGNLTHLPLH
Murine NOGO-R3	LYKCGLSALP	AGVPGGLHSL	QYLYLDQNH	EYLQDDIFVD	LWNLSHLPLH
Human NOGO-R3	LYKCGLSALP	AGVPGGLHSL	QYLYLDQNH	EYLQDDIFVD	LWNLSHLPLH
Human NOGO-R2	LYRCQLSSLP	GNIFRGLVEL	QYLYLDQNSL	LWLQDDLPAD	LWNLSHLPLH
Consensus	L--C-L--L-	---F-GL--L-	QYLYLQ-N--	--L-D--F-D	L--L--HLPLH

(57) Abstract: The invention relates generally to genes that encode proteins that inhibit axonal growth. The invention relates specifically to genes encoding NgR protein homologs in humans and mice. The invention also includes compositions and methods for modulating the expression and activity of Nogo and the NgR proteins. Specifically, the invention includes peptides, proteins and antibodies that block Nogo-mediated inhibition of axonal extension. The compositions and methods of the invention are useful in the treatment of cranial or cerebral trauma, spinal cord injury, stroke or a demyelinating disease.



Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

NOGO RECEPTOR HOMOLOGS

FIELD OF THE INVENTION

5 The invention relates to neurology and molecular biology. More particularly, the invention relates to CNS neurons and axonal growth

BACKGROUND

10 Among the mechanisms through which the cells of an organism communicate with each other and obtain information and stimuli from their environment is through cell membrane receptor molecules expressed on the cell surface. Many such receptors have been identified, characterized, and sometimes classified into major receptor superfamilies based on structural motifs and signal transduction features. The receptors are a first essential link for translating an extracellular signal into a cellular
15 physiological response.

 Receptors on neurons are particularly important in the development of the nervous system during embryogenesis. The neurons form connections with target cells during development through axonal extension of the neurons toward the target cells in a receptor-mediated process. Axons and dendrites have a specialized region of their
20 distal tips known as the growth cone. Growth cones enable the neuron to sense the local environment through a receptor-mediated process and direct the movement of the axon or dendrite of the neuron toward the neuron's target cell. This process is known as elongation. Growth cones can be sensitive to several guidance cues, for example, surface adhesiveness, growth factors, neurotransmitters and electric fields. The
25 guidance of growth at the cone depends on various classes of adhesion molecules, intercellular signals, as well as factors that stimulate and inhibit growth cones.

 Interestingly, damaged neurons do not elongate in the central nervous system (CNS) following injury due to trauma or disease, whereas axons in the peripheral nervous system (PNS) regenerate readily. The fact that damaged CNS neurons fail to
30 elongate is not due to an intrinsic property of CNS axons, but rather due to the CNS environment that is not permissive for axonal elongation. Classical grafting experiments by Aguayo and colleagues (*e.g.*, Richardson *et al.*, (1980) *Nature* 284,

- 2 -

264-265) demonstrated that CNS axons can in fact elongate over substantial distances within peripheral nerve grafts, and that CNS myelin inhibits CNS axon elongation. Therefore, given the appropriate environment, CNS axons can regenerate, implying that CNS axonal injury can potentially be addressed by appropriate manipulation of the
5 CNS environment.

The absence of axon regeneration following injury can be attributed to the presence of axon growth inhibitors. These inhibitors are predominantly associated with myelin and constitute an important barrier to regeneration. Axon growth inhibitors are present in CNS-derived myelin and the plasma membrane of
10 oligodendrocytes that synthesize myelin in the CNS (Schwab *et al.*, (1993) *Annu. Rev. Neurosci.* 16, 565-595). Myelin-associated inhibitors appear to be a primary contributor to the failure of CNS axon regeneration *in vivo* after an interruption of axonal continuity, whereas other non-myelin associated axon growth inhibitors in the CNS may play a lesser role. These inhibitors block axonal regeneration following
15 neuronal injury due to trauma, stroke or viral infection.

Numerous myelin-derived axon growth inhibitors have been characterized (see, for review, David *et al.*, (1999) WO995394547; Bandman *et al.*, (1999) U.S. Patent No. 5,858,708; Schwab, (1996) *Neurochem. Res.* 21, 755-761). Several components of CNS white matter, NI35, NI250 (Nogo) and Myelin-associated glycoprotein
20 (MAG), which have inhibitory activity for axonal extension, have been described as well (Schwab *et al.*, (1990) WO9005191; Schwab *et al.*, (1997) U.S. Patent No. 5,684,133). In particular, Nogo is a 250 kDa myelin-associated axon growth inhibitor that was originally characterized based on the effects of the purified protein *in vitro* and monoclonal antibodies that neutralize the protein's activity (Schwab (1990) *Exp.*
25 *Neurol.* 109, 2-5). The Nogo cDNA was first identified through random analysis of brain cDNA and had no suggested function (Nagase *et al.*, (1998) *DNA Res.* 5, 355-364). The identification of this Nogo cDNA as the cDNA encoding the 250 kDa myelin-associated axon growth inhibitor was discovered only recently (GrandPre *et al.*, (2000) *Nature* 403, 439-444; Chen *et al.*, (2000) *Nature* 403, 434-439; Prinjha *et al.*,
30 (2000) *Nature* 403, 383-384).

Importantly, Nogo has been shown to be the primary component of CNS myelin responsible for inhibiting axonal elongation and regeneration. Nogo's selective

- 3 -

expression by oligodendrocytes and not by Schwann cells (the cells that myelinate P.S. axons) is consistent with the inhibitory effects of CNS myelin, in contrast to P.S. myelin (GrandPre *et al.*, (2000) *Nature* 403, 434-439). In culture, Nogo inhibits axonal elongation and causes growth cone collapse (Spillmann *et al.*, (1998) *J. Biol. Chem.* 272, 19283-19293). Antibodies (e.g., IN-1) against Nogo have been shown to block most of the inhibitory action of CNS myelin on neurite growth *in vitro* (Spillmann *et al.*, (1998) *J. Biol. Chem.* 272:19283-19293). These experiments indicate that Nogo is the main component of CNS myelin responsible for inhibition of axonal elongation in culture. Furthermore, *in vivo*, the IN-1 antibody has been shown to enhance axonal regeneration after spinal cord injury, resulting in recovery of behaviors such as contact placing and stride length (Schnell and Schwab (1990) *Nature* 343, 269-272; Bregman *et al.*, (1995) *Nature* 378, 498-501). Thus, there is substantial evidence that Nogo is a disease-relevant molecular target. Agents that interfere with the binding of Nogo to its receptor would be expected to improve axonal regeneration in clinical states in which axons have been damaged, and improve patient outcome.

Modulation of Nogo has been described as a means for treatment of regeneration for neurons damaged by trauma, infarction and degenerative disorders of the CNS (Schwab *et al.*, (1994) WO9417831; Tatagiba *et al.*, (1997) *Neurosurgery* 40, 541-546) as well as malignant tumors in the CNS such as glioblastoma (Schwab *et al.*, (1993) U.S. Patent No. 5,250,414; Schwab *et al.*, (2000) U.S. Patent No. 6,025,333).

Antibodies which recognize Nogo have been suggested to be useful in the diagnosis and treatment of nerve damage resulting from trauma, infarction and degenerative disorders of the CNS (Schnell & Schwab, (1990) *Nature* 343, 269-272; Schwab *et al.*, (1997) U.S. Patent No. 5,684,133). For CNS axons, there is a correlation between the presence of myelin and the inhibition of axon regeneration over long distances (Savio and Schwab (1990) *Proc. Natl. Acad. Sci.* 87, 4130-4133; Keirstead *et al.*, (1992) *Proc. Natl. Acad. Sci.* 89, 11664-11668). After Nogo is blocked by antibodies, neurons can again extend across lesions caused by nerve damage (Schnell and Schwab (1990) *Nature* 343, 269-272).

- 4 -

SUMMARY OF THE INVENTION

Genes encoding homologs (NgR2 and NgR3) of a Nogo receptor (NgR1) in mice and humans have been discovered. Various domains in the polypeptides encoded by the NgR2 and NgR3 genes have been identified and compared to domains in mouse and human NgR1 polypeptides. This comparison has led to identification of a consensus sequence (NgR consensus sequence) that characterizes a family of proteins (NgR family). Based on these and other discoveries, the invention features molecules and methods for modulating axonal growth in CNS neurons.

The invention provides a polypeptide that contains a polypeptide containing a tryptophan rich LRRCT domain consisting of the amino acid sequence:

N X₁ W X₂ C X₃ C R A R X₄ L W X₅ W X₆ X₇ X₈ X₉ R X₁₀ S S S X₁₁ V

X₁₂ C X₁₃ X₁₄ P X₁₅ X₁₆ X₁₇ X₁₈ X₁₉ X₂₀ D L X₂₁ X₂₂ L X₂₃ X₂₄ X₂₅ D

X₂₆ X₂₇ X₂₈ C [SEQ ID NO: 19]

wherein X is any protein amino acid or a gap, and the polypeptide does not include amino acid sequence from residue 260 to 309 of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1).

Preferably, X17 and X23 are (independently) arginine or lysine. In some embodiments, the amino acid sequence of the LRRCT domain is residues 261-310 of SEQ ID NO:2, or residues 261-310 of SEQ ID NO: 2 with up to 10 conservative amino acid substitutions. In some embodiments, the polypeptide contains the following NTLRRCT amino acid sequence:

C P X₁ X₂ C X₃ C Y X₄ X₅ P X₆ X₇ T X₈ S C X₉ X₁₀ X₁₁ X₁₂ X₁₃ X₁₄ X₁₅ X₁₆ P
X₁₇ X₁₈ X₁₉ P X₂₀ X₂₁ X₂₂ X₂₃ R X₂₄ F L X₂₅ X₂₆ N X₂₇ I X₂₈ X₂₉ X₃₀ X₃₁ X₃₂ X₃₃
X₃₄ F X₃₅ X₃₆ X₃₇ X₃₈ X₃₉ X₄₀ X₄₁ X₄₂ L W X₄₃ X₄₄ S N X₄₅ X₄₆ X₄₇ X₄₈ I X₄₉
X₅₀ X₅₁ X₅₂ F X₅₃ X₅₄ X₅₅ X₅₆ X₅₇ L E X₅₈ L D L X₅₉ D N X₆₀ X₆₁ L X₆₂ X₆₃ X₆₄
X₆₅ P X₆₆ T F X₆₇ G L X₆₈ X₆₉ L X₇₀ X₇₁ L X₇₂ L X₇₃ X₇₄ C X₇₅ L X₇₆ X₇₇ L X₇₈
X₇₉ X₈₀ X₈₁ F X₈₂ G L X₈₃ X₈₄ L Q Y L Y L Q X₈₅ N X₈₆ X₈₇ X₈₈ X₈₉ L X₉₀ D

- 5 -

X₉₁ X₉₂ F X₉₃ D L X₉₄ N L X₉₅ H L F L H G N X₉₆ X₉₇ X₉₈ X₉₉ X₁₀₀ X₁₀₁ X₁₀₂
 X₁₀₃ X₁₀₄ F R G L X₁₀₅ X₁₀₆ L D R L L L H X₁₀₇ N X₁₀₈ X₁₀₉ X₁₁₀ X₁₁₁ V H X₁₁₂
 X₁₁₃ A F X₁₁₄ X₁₁₅ L X₁₁₆ R L X₁₁₇ X₁₁₈ L X₁₁₉ L F X₁₂₀ N X₁₂₁ L X₁₂₂ X₁₂₃ L
 X₁₂₄ X₁₂₅ X₁₂₆ X₁₂₇ L X₁₂₈ X₁₂₉ L X₁₃₀ X₁₃₁ L X₁₃₂ X₁₃₃ L R L N X₁₃₄ N X₁₃₅ W
 5 X₁₃₆ C X₁₃₇ C R X₁₃₈ R X₁₃₉ L W X₁₄₀ W X₁₄₁ X₁₄₂ X₁₄₃ X₁₄₄ R X₁₄₅ S S S X₁₄₆
 V X₁₄₇ C X₁₄₈ X₁₄₉ P X₁₅₀ X₁₅₁ X₁₅₂ X₁₅₃ X₁₅₄ X₁₅₅ D L X₁₅₆ X₁₅₇ L X₁₅₈ X₁₅₉ X₁₆₀
 D X₁₆₁ X₁₆₂ X₁₆₃ C [SEQ ID NO:18]

wherein X is any amino acid residue or a gap and wherein the polypeptide is not the
 10 polypeptide of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1). For
 example, X₆, X₃₇ and X₃₈ may represent a gap. Specific examples of polypeptides of
 the invention are SEQ ID NO: 2 (human NgR2), SEQ ID NO: 4 (mouse NgR3), and
 SEQ ID NO: 14 (human NgR3). In some embodiments, the polypeptide contains: (a)
 a NTLRRCT domain, and (b) less than a complete CTS domain, provided that a partial
 15 CTS domain, if present, consists of no more than the first 39 amino acids of the CTS
 domain. While the polypeptide may contain a functional GPI domain, a functional GPI
 domain may be absent, e.g., when a soluble polypeptide is desired. A polypeptide of
 the invention optionally includes an amino acid sequence of a heterologous
 polypeptide, e.g., an Fc portion of an antibody.

20 The invention also provides a nucleic acid encoding an above-described
 polypeptide; a vector containing the nucleic acid, which nucleic acid may be operably
 linked to an expression control sequence; and a transformed host cell containing the
 vector. A method of producing a polypeptide of the invention is also provided. The
 method includes introducing a nucleic acid encoding the above-described polypeptide
 25 into a host cell, culturing the cell under conditions suitable for expression of the
 polypeptide, and recovering the polypeptide.

The invention also provides an antisense molecule whose nucleotide sequence
 is complementary to a nucleotide sequence encoding a polypeptide selected from the
 group consisting of: a polypeptide consisting of residues 311-395 of SEQ ID NO: 2, a
 30 polypeptide consisting of residues 256-396 of SEQ ID NO:14 and a polypeptide
 consisting of residues 321-438 of SEQ ID NO: 4, wherein the nucleic acid is from 8 to

- 6 -

100 nucleotides in length, e.g., about 20, 30, 40, 50, 60, 70, 80 or 90 nucleotides. The invention also provides a nucleic acid encoding such an antisense molecule.

The invention also provides an antibody that binds to an above-described polypeptide. Polypeptides or antibodies of the invention can be formulated into
5 pharmaceutical compositions containing the polypeptide or antibody and a pharmaceutically acceptable carrier.

The invention also provides a method for decreasing inhibition of axonal growth of a CNS neuron. The method includes the step of contacting the neuron with an effective amount of a polypeptide or antibody of the invention.

10 The invention also provides a method for treating a central nervous system disease, disorder or injury. The method includes administering to a mammal, e.g., a human, an effective amount of a polypeptide or antibody of the invention. Exemplary diseases, disorders and injuries that may be treated using molecules and methods of the invention include, but are not limited to, cerebral injury, spinal cord injury, stroke,
15 demyelinating diseases, e.g., multiple sclerosis, monophasic demyelination, encephalomyelitis, multifocal leukoencephalopathy, panencephalitis, Marchiafava-Bignami disease, Spongy degeneration, Alexander's disease, Canavan's disease, metachromatic leukodystrophy and Krabbe's disease.

The invention also provides a method for identifying a molecule that binds a
20 polypeptide of the invention. The method includes the steps of: (a) providing a polypeptide of the invention; (b) contacting the polypeptide with the candidate molecule; and (c) detecting binding of the candidate molecule to the polypeptide.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the
25 invention belongs. In case of conflict, the present application, including definitions, will control. All publications, patent and other references mentioned herein are incorporated by reference.

The materials, methods and examples presented below are illustrative only, and not intended to be limiting. Other features and advantages of the invention will be
30 apparent from the detail description and from the claims.

BRIEF DESCRIPTION OF THE FIGURES

- 7 -

Fig. 1A-1B shows an alignment of NgR2 (SEQ ID NO:2) and NgR3 (SEQ ID NO:4) with the known NgR, NgR1 (SEQ ID NO:5) and the Consensus Sequence (SEQ ID NO:6).

5

Fig. 2. mNgR3 does not bind hNogoA(1055-1120). COS-7 cells were transfected with vectors encoding myc-NgR1 or myc-NgR3, fixed, and stained with anti-myc antibodies or AP-hNogoA(1055-1120).

10

Fig.3. An alignment of the amino acid sequences of human NgR1, murine NgR1, murine NgR3, human NgR3 and human NgR2. Numbering begins with amino acid #1 of murine NgR3. The consensus sequence is listed below. The LRR NT domain is indicated by a shaded box; domains LLR 1, LLR 3, LLR 5, and LLR 7 are indicated by open boxes; LLR 2, LLR 4, LLR 6 and LLR 8 are indicated by shaded boxes; and the LLR CT domain is indicated by a shaded box. Amino acids in bold in LLR 8 indicate a conserved glycosylation sites. A dot indicates conserved cystine residue in LLR4. Box at C terminus indicates putative GPI signals.

15

20

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides purified and isolated polynucleotides (*e.g.*, DNA sequences and RNA transcripts, both sense and complementary antisense strands, both single- and double-stranded, including splice variants thereof) encoding NgR homologs, referred to herein as NgR. Unless indicated otherwise, as used herein, the abbreviation in lower case (NgR) refers to a gene, cDNA, RNA or nucleic acid sequence, whereas the upper case version (NgR) refers to a protein, polypeptide, peptide, oligopeptide, or amino acid sequence. Specific proteins are designated by number, *e.g.*, "NgR2" is a human NgR homolog, "NgR3" is a murine-derived NgR homolog, and "NgR1" is the known NgR identified by Dr. Stephen Strittmatter.

Known NgRs are herein referred to as "NgRs." DNA polynucleotides of the invention include genomic DNA, cDNA and DNA that has been chemically synthesized in whole or in part.

25

30

- 8 -

Standard reference works setting forth the general principles of recombinant DNA technology known to those of skill in the art include Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York (1998); Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d Ed., Cold Spring Harbor Laboratory Press, Plainview, New York (1989); Kaufman *et al.*, Eds., HANDBOOK OF MOLECULAR AND CELLULAR METHODS IN BIOLOGY AND MEDICINE, CRC Press, Boca Raton (1995); McPherson, Ed., DIRECTED MUTAGENESIS: A PRACTICAL APPROACH, IRL Press, Oxford (1991).

As used herein, the term "axon" refers to a long cellular protrusion from a neuron, whereby action potentials are conducted, either to or from the cell body.

As used herein, the term "axonal growth" refers to an extension of the long process or axon, originating at the cell body and proceeded by the growth cone.

As used herein, the term "central nervous system disorder" refers to any pathological state associated with abnormal function of the central nervous system (CNS). The term includes, but is not limited to, altered CNS function resulting from physical trauma to cerebral tissue, viral infection, autoimmune mechanisms and genetic mutation.

As used herein, the term "demyelinating disease" refers to a pathological disorder characterized by the degradation of the myelin sheath of the oligodendrocyte cell membrane.

As used herein, the term "growth cone" refers to a specialized region at the tip of a growing neurite that is responsible for sensing the local environment and moving the axon toward its appropriate synaptic target cell.

As used herein, the term "growth cone movement" refers to the extension or collapse of the growth cone toward a neuron's target cell.

As used herein, the term "neurite" refers to a process growing out of a neuron. As it is sometimes difficult to distinguish a dendrite from an axon in culture, the term "neurite" is used for both.

As used herein, the term "oligodendrocyte" refers to a neuroglial cell of the CNS whose function is to myelinate CNS axons.

"Synthesized" as used herein and understood in the art, refers to polynucleotides produced by purely chemical, as opposed to enzymatic, methods.

- 9 -

"Wholly" synthesized DNA sequences are therefore produced entirely by chemical means, and "partially" synthesized DNAs embrace those wherein only portions of the resulting DNA were produced by chemical means. By the term "region" is meant a physically contiguous portion of the primary structure of a biomolecule. In the case of proteins, a region is defined by a contiguous portion of the amino acid sequence of that protein. The term "domain" is herein defined as referring to a structural part of a biomolecule that contributes to a known or suspected function of the biomolecule. Domains may be co-extensive with regions or portions thereof; domains may also incorporate a portion of a biomolecule that is distinct from a particular region, in addition to all or part of that region. Examples of NgR protein domains include, but are not limited to, the signal peptide, extracellular (*i.e.*, N-terminal) domain, and leucine-rich repeat domains.

As used herein, the term "activity" refers to a variety of measurable indicia suggesting or revealing binding, either direct or indirect; affecting a response, *i.e.*, having a measurable affect in response to some exposure or stimulus, including, for example, the affinity of a compound for directly binding a polypeptide or polynucleotide of the invention, or, for example, measurement of amounts of upstream or downstream proteins or other similar functions after some stimulus or event. Such activities may be measured by assays such as competitive inhibition of NgR1 binding to Nogo assays wherein, for example, unlabeled, soluble NgR2 is added to an assay system in increasing concentrations to inhibit the binding of Nogo to NgR1 expressed on the surface of CHO cells. As another example, one may assess the ability of neurons to extend across lesions caused by nerve damage (as in Schnell and Schwab (1990) *Nature* 343, 269-272) following inhibition of Nogo by various forms of NgR2 and/or NgR3 as a biological indicator of NgR function.

As used herein, the term "antibody" is meant to refer to complete, intact antibodies, and Fab, Fab', F(ab)2, and other fragments thereof. Complete, intact antibodies include monoclonal antibodies such as murine monoclonal antibodies, chimeric antibodies, anti-idiotypic antibodies, anti-anti-idiotypic antibodies, and humanized antibodies.

As used herein, the term "binding" means the physical or chemical interaction between two proteins or compounds or associated proteins or compounds or

- 10 -

combinations thereof. Binding includes ionic, non-ionic, hydrogen bonds, Van der Waals, hydrophobic interactions, etc. The physical interaction, the binding, can be either direct or indirect, indirect being through or due to the effects of another protein or compound. Direct binding refers to interactions that do not take place through or
5 due to the effect of another protein or compound but instead are without other substantial chemical intermediates.

As used herein, the term "compound" means any identifiable chemical or molecule, including, but not limited to, small molecules, peptides, proteins, sugars, nucleotides or nucleic acids, and such compound can be natural or synthetic.

10 As used herein, the term "complementary" refers to Watson-Crick basepairing between nucleotide units of a nucleic acid molecule.

As used herein, the term "contacting" means bringing together, either directly or indirectly, a compound into physical proximity to a polypeptide or polynucleotide of the invention. The polypeptide or polynucleotide can be in any number of buffers,
15 salts, solutions etc. Contacting includes, for example, placing the compound into a beaker, microtiter plate, cell culture flask, or a microarray, such as a gene chip, or the like, which contains the nucleic acid molecule, or polypeptide encoding the NgR or fragment thereof.

As used herein, the phrase "homologous nucleotide sequence," or "homologous
20 amino acid sequence," or variations thereof, refers to sequences characterized by an identity at the nucleotide level, or a homology at the amino acid level, of at least the specified percentage. Homologous nucleotide sequences include those sequences coding for isoforms of proteins. Such isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA.

25 Alternatively, isoforms can be encoded by different genes. Homologous nucleotide sequences include nucleotide sequences encoding for a protein of a species other than humans, including, but not limited to, mammals. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does
30 not, however, include the nucleotide sequence encoding NgR1. Homologous amino acid sequences include those amino acid sequences which contain conservative amino acid substitutions and which polypeptides have the same binding and/or activity. A

- 11 -

homologous amino acid sequence does not, however, include the amino acid sequence encoding other known NgRs. Percent homology can be determined by, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison WI), using the default
5 settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489, which is incorporated herein by reference in its entirety).

As used herein, the term "isolated" nucleic acid molecule refers to a nucleic acid molecule (DNA or RNA) that is substantially free of nucleic acids encoding other proteins with which it is associated in nature, i.e., a nucleic acid that has been removed
10 from its native environment. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*,
15 sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated NgR nucleic acid molecule can contain less than about 50 kb, 25 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the
20 nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

As used herein, the term "heterologous" refers to a nucleotide or amino acid
25 sequence that is a different, or non-corresponding sequence, or a sequence derived from a different species. For example, a mouse NgR nucleotide or amino acid sequence is heterologous to a human NgR nucleotide or amino acid sequence, and a human NgR nucleic or amino acid sequence is heterologous to a human immunoglobulin nucleotide or amino acid sequence.

30 As used herein, a "soluble NgR polypeptide" is a NgR polypeptide that does not anchor itself in a membrane. Such soluble polypeptides include, for example, NgR2 and NgR3 polypeptides that lack a sufficient portion of their GPI anchor signal

to anchor the polypeptide or are modified such that the GPI anchor signal is not adequate to result in replacement of the peptide with a GPI anchor. In preferred embodiments, up to 5, 10, 20 or 25 amino acids are removed from the C-terminus of NgR2 or NgR3 to make the respective proteins soluble. As used herein soluble NgR polypeptides include full-length or truncated (*e.g.*, with internal deletions) NgR.

Soluble NgR polypeptides may include the entire NgR protein up to the putative GPI signal sequence (*e.g.*, amino acid 1 to about amino acid 395 of NgR2, and from amino acid 1 to about amino acid 438 of NgR3). In other embodiments, the signal peptide of the proteins may be removed or truncated (*e.g.*, all or part of the signal sequence of NgR2, which spans amino acid 1 to about amino acid 30 of SEQ ID NO:2, may be removed; all or part of the signal sequence of NgR3, which spans amino acid 1 to about amino acid 40 of SEQ ID NO:4, may be removed). In some embodiments, the mature NgR2 (SEQ ID NO:8) and the mature NgR3 (SEQ ID NO:9) are used.

Soluble NgR polypeptides include at least one of the putative ligand-binding portions of NgR, including the first cysteine-rich region (SEQ ID NO:10, the leucine repeat region (SEQ ID NO:12) and the second cysteine-rich region (SEQ ID NO:11). In some embodiments, soluble NgR polypeptides consist of amino acid 1 through about amino acid 395 of SEQ ID NO:2, or amino acid 1 through about amino acid 438 of SEQ ID NO:4.

In other embodiments, the soluble NgR polypeptides are fusion proteins that contain amino acids 30 through about amino acid 395 of mature NgR2 or amino acid 40 through about amino acid 438 of NgR3, the C-terminal 10 amino acids of a human IgG 1 hinge region containing the two cysteine residues thought to participate in interchain disulfide bonding, and the CH2 and CH3 regions of a human IgG1 heavy chain constant domain. This type of recombinant protein is designed to modulate inhibition of axonal elongation through inhibition of the Nogo ligand binding to NgR1, or by inhibiting the ligand of the NgR from interacting with cell surface NgR. The NgR portion of the fusion binds to the Nogo ligand and the IgG1 portion binds to the FcγRI (macrophage) and FcγIII (NK cells and neutrophils) receptors.

The production of the soluble polypeptides useful in this invention may be achieved by a variety of methods known in the art. For example, the polypeptides may

- 13 -

be derived from intact transmembrane NgR molecules by proteolysis using specific endopeptidases in combination with exopeptidases, Edman degradation, or both. The intact NgR molecule, in turn, may be purified from its natural source using conventional methods. Alternatively, the intact NgR may be produced by known
5 recombinant DNA techniques using cDNAs, expression vectors and well-known techniques for recombinant gene expression.

Preferably, the soluble polypeptides useful in the present invention are produced directly, thus eliminating the need for an entire NgR as a starting material. This may be achieved by conventional chemical synthesis techniques or by well-known
10 recombinant DNA techniques wherein only those DNA sequences which encode the desired peptides are expressed in transformed hosts. For example, a gene which encodes the desired soluble NgR polypeptide may be synthesized by chemical means using an oligonucleotide synthesizer. Such oligonucleotides are designed based on the amino acid sequence of the desired soluble NgR polypeptide. Specific DNA sequences
15 coding for the desired peptide also can be derived from the full-length DNA sequence by isolation of specific restriction endonuclease fragments or by PCR synthesis of the specified region from cDNA.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NOs:1, 3 or a complement of either of
20 these nucleotide sequences, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequences of SEQ ID NOs:1 or 3 as a hybridization probe, NgR nucleic acid sequences can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, eds., MOLECULAR CLONING: A LABORATORY
25 MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993).

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers
30 according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis.

- 14 -

Furthermore, oligonucleotides corresponding to NgR nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the terms "modulates" or "modifies" means an increase or decrease in the amount, quality, or effect of a particular activity or protein.

5 As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues which has a sufficient number of bases to be used in a polymerase chain reaction (PCR). This short sequence is based on (or designed from) a genomic or cDNA sequence and is used to amplify, confirm or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue.

10 Oligonucleotides comprise portions of a DNA sequence having at least about 10 nucleotides and as many as about 50 nucleotides, preferably about 15 to 30 nucleotides. They are chemically synthesized and may be used as probes.

 As used herein, the term "probe" refers to nucleic acid sequences of variable length, preferably between at least about 10 and as many as about 6,000 nucleotides,
15 depending on use. They are used in the detection of identical, similar or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. They may be single- or double-stranded and carefully designed to have specificity in PCR, hybridization membrane-based, or ELISA-like technologies.

20 The term "preventing" refers to decreasing the probability that an organism contracts or develops an abnormal condition.

 The term "treating" refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

 The term "therapeutic effect" refers to the inhibition or activation factors
25 causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (*i.e.*, slowing or stopping) of cell death; (c) inhibition of
30 degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population

- 15 -

of cells. Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

The term "abnormal condition" refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal
5 condition can relate to cell proliferation, cell differentiation, cell signaling, or cell survival. An abnormal condition may also include obesity, diabetic complications such as retinal degeneration, and irregularities in glucose uptake and metabolism, and fatty acid uptake and metabolism.

Abnormal cell proliferative conditions, for example, include cancers such as
10 fibrotic and mesangial disorders, abnormal angiogenesis and vasculogenesis, wound healing, psoriasis, diabetes mellitus and inflammation.

Abnormal differentiation conditions include, for example, neurodegenerative disorders, slow wound healing rates and slow tissue grafting healing rates.

Abnormal cell signaling conditions include, for example, psychiatric disorders
15 involving excess neurotransmitter activity.

Abnormal cell survival conditions may also relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the function of any one of the protein kinases could lead to cell immortality or premature
20 cell death.

The term "administering" relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell
25 culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques and carrier techniques.

30 The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an organism. The effect of administering a compound on organism function can then

- 16 -

be monitored. The organism is preferably a mouse, rat, rabbit, guinea pig or goat, more preferably a monkey or ape, and most preferably a human.

By "amplification" it is meant increased numbers of DNA or RNA in a cell compared with normal cells. "Amplification" as it refers to RNA can be the detectable
5 presence of RNA in cells, since in some normal cells there is no basal expression of RNA. In other normal cells, a basal level of expression exists, therefore in these cases amplification is the detection of at least 1-2-fold, and preferably more, compared to the basal level.

The amino acid sequences are presented in the amino to carboxy direction,
10 from left to right. The amino and carboxy groups are not presented in the sequence. The nucleotide sequences are presented by single strand only, in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission or (for amino acids) by three letters code.

15

Nucleic Acids

Genomic DNA of the invention comprises the protein-coding region for a polypeptide of the invention and is also intended to include allelic variants thereof. It is widely understood that, for many genes, genomic DNA is transcribed into RNA
20 transcripts that undergo one or more splicing events wherein intron (*i.e.*, non-coding regions) of the transcripts are removed, or "spliced out." RNA transcripts that can be spliced by alternative mechanisms, and therefore be subject to removal of different RNA sequences but still encode a NgR polypeptide, are referred to in the art as splice variants which are embraced by the invention. Splice variants comprehended by the
25 invention therefore are encoded by the same original genomic DNA sequences but arise from distinct mRNA transcripts. Allelic variants are modified forms of a wild-type gene sequence, the modification resulting from recombination during chromosomal segregation or exposure to conditions which give rise to genetic mutation. Allelic variants, like wild-type genes, are naturally occurring sequences (as
30 opposed to non-naturally occurring variants arising from in vitro manipulation).

The invention also comprehends cDNA that is obtained through reverse transcription of an RNA polynucleotide encoding NgR (conventionally followed by

- 17 -

second-strand synthesis of a complementary strand to provide a double-stranded DNA).

Preferred DNA sequences encoding a human NgR polypeptide is set out in SEQ ID NOs:1 and 13. A preferred DNA of the invention comprises a double
5 stranded molecule comprising the coding molecule (*i.e.*, the "coding strand") along with the complementary molecule (the "non-coding strand" or "complement") having a sequence unambiguously deducible from the coding strand according to Watson-Crick base-pairing rules for DNA. Also preferred are other polynucleotides encoding NgR polypeptides, as shown in SEQ ID NO:3, which comprises murine NgR homolog,
10 NgR3.

Also preferred are nucleotide sequences that encode at least a portion of a NgR polypeptide that has at least one biological function of a NgR. More preferred are nucleotide sequences that encode a portion of NgR that encodes at least the mature NgR without the hydrophobic C-terminal GPI signal. Also preferred are nucleotide
15 sequences that encode the portion of NgR that encodes at least the ligand-binding region of NgR.

The invention further embraces other species, preferably mammalian, homologs of the human NgR DNA. Species homologs, sometimes referred to as "orthologs," in general, share at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at
20 least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% homology with human DNA of the invention. Generally, percent sequence "homology" with respect to polynucleotides of the invention may be calculated as the percentage of nucleotide bases in the candidate sequence that are identical to nucleotides in the NgR sequences set forth in SEQ ID
25 NOs:1, 3 or 13, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity.

The polynucleotide sequence information provided by the invention makes possible large-scale expression of the encoded polypeptide by techniques well known and routinely practiced in the art. Polynucleotides of the invention also permit
30 identification and isolation of polynucleotides encoding related NgR polypeptides, such as human allelic variants and species homologs, by well-known techniques including Southern and/or Northern hybridization, and polymerase chain reaction (PCR).

- 18 -

Examples of related polynucleotides include human and non-human genomic sequences, including allelic variants, as well as polynucleotides encoding polypeptides homologous to NgR and structurally related polypeptides sharing one or more biological, immunological, and/or physical properties of NgR. Non-human species genes encoding proteins homologous to NgR can also be identified by Southern and/or PCR analysis and are useful in animal models for NgR disorders. Knowledge of the sequence of a human NgR DNA also makes possible through use of Southern hybridization or polymerase chain reaction (PCR) the identification of genomic DNA sequences encoding NgR expression control regulatory sequences such as promoters, operators, enhancers, repressors, and the like. Polynucleotides of the invention are also useful in hybridization assays to detect the capacity of cells to express NgR. Polynucleotides of the invention may also provide a basis for diagnostic methods useful for identifying a genetic alteration(s) in a NgR locus that underlies a disease state or states, which information is useful both for diagnosis and for selection of therapeutic strategies.

The disclosure herein of a full-length polynucleotide encoding a NgR polypeptide makes readily available to the worker of ordinary skill in the art every possible fragment of the full-length polynucleotide. The invention, therefore, provides fragments of NgR-encoding polynucleotides comprising at least 6, and preferably at least 14, 16, 18, 20, 25, 50, or 75 consecutive nucleotides of a polynucleotide encoding NgR. Preferably, fragments of polynucleotides of the invention comprise sequences unique to the NgR-encoding polynucleotide sequence, and therefore hybridize under highly stringent or moderately stringent conditions only (*i.e.*, "specifically") to polynucleotides encoding NgR (or fragments thereof).

Polynucleotide fragments of genomic sequences of the invention comprise not only sequences unique to the coding region, but also include fragments of the full-length sequence derived from introns, regulatory regions, and/or other non-translated sequences. Sequences unique to polynucleotides of the invention are recognizable through sequence comparison to other known polynucleotides, and can be identified through use of alignment programs routinely utilized in the art, *e.g.*, those made available in public sequence databases. Such sequences also are recognizable from Southern hybridization analyses to determine the number of fragments of genomic

- 19 -

DNA to which a polynucleotide will hybridize. Polynucleotides of the invention can be labeled in a manner that permits their detection, including radioactive, fluorescent and enzymatic labeling.

5 Fragments of polynucleotides are particularly useful as probes for detection of full-length or fragment of NgR polynucleotides. One or more polynucleotides can be included in kits that are used to detect the presence of a polynucleotide encoding NgR, or used to detect variations in a polynucleotide sequence encoding NgR.

 The invention also embraces DNAs encoding NgR polypeptides that hybridize under moderately stringent or high stringency conditions to the noncoding strand, or
10 complement, of the polynucleotide in any of SEQ ID NOs:1 or 3.

 Stringent conditions are known to those skilled in the art and can be found in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98% or 99% homologous to each other typically remain
15 hybridized to each other. A non-limiting example of stringent hybridization conditions is hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA and 500 mg/ml denatured salmon sperm DNA at 65°C. This hybridization is followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that
20 hybridizes under stringent conditions to the sequence of SEQ ID NOs:1 or 3 corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein). As used herein, "stringent hybridization conditions" means: 42°C in a hybridization
25 solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% (wt/vol) dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS.

Vectors

30 Another aspect of the present invention is directed to vectors, or recombinant expression vectors, comprising any of the nucleic acid molecules described above. Vectors are used herein either to amplify DNA or RNA encoding NgR and/or to

- 20 -

express DNA which encodes NgR. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of
5 vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction
10 into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used
15 interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), that serve equivalent functions.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with
20 vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (1) to increase expression of recombinant protein; (2) to increase the solubility of the recombinant protein; and (3)
25 to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition
30 sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67, 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia,

- 21 -

Piscataway, N.J.) that fuse glutathione-S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69, 301-315) and pET 11d (Studier *et al.*, GENE
5 EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, Gottesman, GENE EXPRESSION TECHNOLOGY:
10 METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20, 2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by
15 standard DNA synthesis techniques.

In another embodiment, the NgR expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *EMBO J.* 6, 229-234), pMFa (Kurjan and Herskowitz (1982) *Cell* 30, 933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54, 113-123), pYES2
20 (Invitrogen Corporation, San Diego, CA), and picZ (Invitrogen Corp, San Diego, CA).

Alternatively, NgR can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, SF9 cells) include the pAc series (Smith *et al.*, (1983) *Mol. Cell. Biol.* 3,
25 2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170, 31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329, 840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6, 187-195). When used in mammalian cells, the
30 expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both

- 22 -

prokaryotic and eukaryotic cells. See, *e.g.*, Chapters 16 and 17 of Sambrook *et al.*, (Eds.) MOLECULAR CLONING: A LABORATORY MANUAL. 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is
5 capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1, 268-277), lymphoid-specific promoters (Calame and Eaton
10 (1988) *Adv. Immunol.* 43, 235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8, 729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33, 729-740; Queen and Baltimore (1983) *Cell* 33, 741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86, 5473-5477), pancreas-specific promoters (Edlund *et al.*
15 (1985) *Science* 230, 912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, *e.g.*, the murine hox promoters (Kessel and Gruss (1990) *Science* 249, 374-379) and the α -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3, 537-546).

20 The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense NgR mRNA. Regulatory sequences operatively linked
25 to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue-specific or cell-type-specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid,
30 phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation

- 23 -

of gene expression using antisense genes see Weintraub *et al.*, *Antisense RNA as a molecular tool for genetic analysis*, REVIEWS--TRENDS IN GENETICS, Vol. 1(1) 1986.

Preferred vectors include, but are not limited to, plasmids, phages, cosmids, episomes, viral particles or viruses and integratable DNA fragments (*i.e.*, fragments integratable into the host genome by homologous recombination). Preferred viral particles include, but are not limited to, adenoviruses, baculoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses and retroviruses. Preferred expression vectors include, but are not limited to, pcDNA3 (Invitrogen) and pSVL (Pharmacia Biotech). Other expression vectors include, but are not limited to, pSPORT™ vectors, pGEM™ vectors (Promega), pPROEXvectors™ (LTI, Bethesda, MD), Bluescript™ vectors (Stratagene), pQE™ vectors (Qiagen), pSE420™ (Invitrogen) and pYES2™ (Invitrogen).

Preferred expression vectors are replicable DNA constructs in which a DNA sequence encoding NgR is operably linked or connected to suitable control sequences capable of effecting the expression of the NgR in a suitable host. DNA regions are operably linked or connected when they are functionally related to each other. For example, a promoter is operably linked or connected to a coding sequence if it controls the transcription of the sequence. Amplification vectors do not require expression control domains, but rather need only the ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants. The need for control sequences in the expression vector will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include, but are not limited to a transcriptional promoter, enhancers, an optional operator sequence to control transcription, polyadenylation signals, a sequence encoding suitable mRNA ribosomal binding and sequences which control the termination of transcription and translation. Such regulatory sequences are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as

- 24 -

the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, NgR proteins, mutant forms of NgR, fusion
5 proteins, etc.).

Preferred vectors preferably contain a promoter that is recognized by the host organism. The promoter sequences of the present invention may be prokaryotic, eukaryotic or viral. Examples of suitable prokaryotic sequences include the PR and PL promoters of bacteriophage lambda (THE BACTERIOPHAGE LAMBDA, Hershey, A.D.
10 (Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1973), which is incorporated herein by reference in its entirety; LAMBDA II, Hendrix, R.W. (Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1980), which is incorporated herein by reference in its entirety); the *trp*, *recA*, heat shock, and *lacZ* promoters of *E. coli* and the SV40 early promoter (Benoist *et al.*, (1981) *Nature* 290,
15 304-310, which is incorporated herein by reference in its entirety). Additional promoters include, but are not limited to, mouse mammary tumor virus, long terminal repeat of human immunodeficiency virus, maloney virus, cytomegalovirus immediate early promoter, Epstein Barr virus, Rous sarcoma virus, human actin, human myosin, human hemoglobin, human muscle creatine and human metallothionein.

20 Additional regulatory sequences can also be included in preferred vectors. Preferred examples of suitable regulatory sequences are represented by the Shine-Dalgarno sequence of the replicase gene of the phage MS-2 and of the gene *cII* of bacteriophage lambda. The Shine-Dalgarno sequence may be directly followed by DNA encoding NgR and result in the expression of the mature NgR protein.

25 Moreover, suitable expression vectors can include an appropriate marker that allows the screening of the transformed host cells. The transformation of the selected host is carried out using any one of the various techniques well known to the expert in the art and described in Sambrook *et al.*, *supra*.

An origin of replication can also be provided either by construction of the
30 vector to include an exogenous origin or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter may be sufficient. Alternatively, rather than using vectors

- 25 -

which contain viral origins of replication, one skilled in the art can transform mammalian cells by the method of co-transformation with a selectable marker and NgR DNA. An example of a suitable marker is dihydrofolate reductase (DHFR) or thymidine kinase (*see*, U.S. Patent No. 4,399,216).

5 Nucleotide sequences encoding NgR may be recombined with vector DNA in accordance with conventional techniques, including blunt-ended or staggered-ended termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining and ligation with appropriate ligases. Techniques for such manipulation are disclosed by Sambrook *et al.*, *supra* and are well known in the art. Methods for
10 construction of mammalian expression vectors are disclosed in, for example, Okayama *et al.*, (1983) *Mol. Cell. Biol.* 3:280, Cosman *et al.* (1986) *Mol. Immunol.* 23:935, Cosman *et al.*, (1984) *Nature* 312:768, EP-A-0367566, and WO 91/18982, each of which is incorporated herein by reference in its entirety.

15

Host Cells and Transformed Host Cells

According to another aspect of the invention, host cells are provided, including prokaryotic and eukaryotic cells, comprising a polynucleotide of the invention (or vector of the invention) in a manner that permits expression of the encoded NgR
20 polypeptide. Preferably, the cell produces little or no endogenous NgR polypeptide. Polynucleotides of the invention may be introduced into the host cell as part of a circular plasmid, or as linear DNA comprising an isolated protein coding region or a viral vector. Methods for introducing DNA into the host cell that are well known and routinely practiced in the art include transformation, transfection, electroporation,
25 nuclear injection, or fusion with carriers such as liposomes, micelles, ghost cells and protoplasts. Expression systems of the invention include bacterial, yeast, fungal, plant, insect, invertebrate, vertebrate and mammalian cells systems.

Host cells of the invention are a valuable source of immunogen for development of antibodies specifically immunoreactive with NgR. Host cells of the
30 invention are also useful in methods for the large-scale production of NgR polypeptides wherein the cells are grown in a suitable culture medium and the desired polypeptide products are isolated from the cells, or from the medium in which the cells

- 26 -

are grown, by purification methods known in the art, *e.g.*, conventional chromatographic methods including immunoaffinity chromatography, receptor affinity chromatography, hydrophobic interaction chromatography, lectin affinity chromatography, size exclusion filtration, cation or anion exchange chromatography, high pressure liquid chromatography (HPLC), reverse phase HPLC, and the like. Still other methods of purification include those methods wherein the desired protein is expressed and purified as a fusion protein having a specific tag, label or chelating moiety that is recognized by a specific binding partner or agent. The purified protein can be cleaved to yield the desired protein, or can be left as an intact fusion protein.

10 Cleavage of the fusion component may produce a form of the desired protein having additional amino acid residues as a result of the cleavage process.

Knowledge of NgR DNA sequences allows for modification of cells to permit, or increase, expression of endogenous NgR. Cells can be modified (*e.g.*, by homologous recombination) to provide increased expression by replacing, in whole or

15 in part, the naturally occurring NgR promoter with all or part of a heterologous promoter so that the cells express NgR at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to endogenous NgR encoding sequences. (See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No.

20 WO 91/09955.) It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, *ada*, *dhfr*, and the multifunctional CAD gene which encodes carbamoyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the NgR coding sequence, amplification of the marker

25 DNA by standard selection methods results in co-amplification of the NgR coding sequences in the cells.

The DNA sequence information provided by the present invention also makes possible the development (*e.g.*, by homologous recombination or "knock-out" strategies; see Capecchi, *Science* 244:1288-1292 (1989)) of animals that fail to express

30 functional NgR or that express a variant of NgR. Such animals (especially small laboratory animals such as rats, rabbits and mice) are useful as models for studying the *in vivo* activities of NgR and modulators of NgR.

- 27 -

Suitable host cells for expression of the polypeptides of the invention include, but are not limited to, prokaryotes, yeast, and eukaryotes. If a prokaryotic expression vector is employed, then the appropriate host cell would be any prokaryotic cell capable of expressing the cloned sequences. Suitable prokaryotic cells include, but are not limited to, bacteria of the genera *Escherichia*, *Bacillus*, *Salmonella*, *Pseudomonas*,
5 *Streptomyces* and *Staphylococcus*.

If a eukaryotic expression vector is employed, then the appropriate host cell would be any eukaryotic cell capable of expressing the cloned sequence. Preferably, eukaryotic cells are cells of higher eukaryotes. Suitable eukaryotic cells include, but are not limited to, non-human mammalian tissue culture cells and human tissue culture
10 cells. Preferred host cells include, but are not limited to, insect cells, HeLa cells, Chinese hamster ovary cells (CHO cells), African green monkey kidney cells (COS cells), human 293 cells, and murine 3T3 fibroblasts. Propagation of such cells in cell culture has become a routine procedure (*see*, Tissue Culture, Academic Press, Kruse
15 and Patterson, Eds. (1973), which is incorporated herein by reference in its entirety).

In addition, a yeast cell may be employed as a host cell. Preferred yeast cells include, but are not limited to, the genera *Saccharomyces*, *Pichia* and *Kluveromyces*. Preferred yeast hosts are *S. cerevisiae* and *P. pastoris*. Preferred yeast vectors can contain an origin of replication sequence from a 2T yeast plasmid, an autonomously
20 replication sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination and a selectable marker gene. Shuttle vectors for replication in both yeast and *E. coli* are also included herein.

Alternatively, insect cells may be used as host cells. In a preferred embodiment, the polypeptides of the invention are expressed using a baculovirus
25 expression system (*see*, Luckow *et al.*, *Bio/Technology*, 1988, 6, 47; BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL, O'Rielly *et al.* (Eds.), W.H. Freeman and Company, New York, 1992; and U.S. Patent No. 4,879,236, each of which is incorporated herein by reference in its entirety). In addition, the MAXBAC™ complete baculovirus expression system (Invitrogen) can, for example, be used for
30 production in insect cells.

Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA

- 28 -

(1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via
5 conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or
10 transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells
15 may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin, dihydrofolate reductase (DHFR) and methotrexate. Nucleic acid
20 encoding a selectable marker can be introduced into a host cell on the same vector as that encoding NgR or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

In a preferred embodiment, the polypeptides of the invention, including forms
25 of NgR2 and NgR3, soluble forms of NgR, chimeric NgR polypeptides, NgR/Ig fusions and fragments and variations of each of the above are expressed in Chinese Hamster Ovary (CHO) cells.

In order to introduce the DNA fragment coding for the NgR protein or polypeptide into the CHO cell to express the recombinant NgR protein or polypeptide,
30 it is necessary to construct the expression vector.

The vectors for CHO expression include, but are not limited to, pA1-11, pXT1, pRc/CMV, pRc/RSV and pcDNAINeo. The promoter is not specifically limited

- 29 -

provided it effectively promotes expression in CHO cells. Examples of suitable promoters are: SR α , SV40, LTR, CMV, and HSV-TK. Of these, CMV and SR α promoters are preferred.

5 In addition to the above-mentioned promoters, the expression vectors may contain enhancers, splicing signals, polyadenylation signals, selectable markers and an SV40 replication origin. Suitable selectable markers include, but are not limited to the dihydrofolate reductase (DHFR) gene which provides resistance to methotrexate (MTX), the ampicillin resistance gene, and the neomycin resistance gene.

10 Examples of the expression vectors each containing the DNA coding for NgR, portions, fragments and soluble constructs thereof, include the vector (such as one described above), into which the promoter is operably linked (preferably upstream) to the nucleotide sequence encoding the desired NgR construct; a polyadenylation signal downstream from the nucleotide sequence encoding the NgR construct; and, preferably, the vector includes an operable DHFR gene. Preferably, the ampicillin
15 resistant gene is also operably contained in the vector.

CHO cell lacking the DHFR gene (Urlaub, G. *et al.*, (1980) *Proc. Natl. Acad. Sci. USA* 77, 4216-4220) and CHO-K1 (*Proc. Natl. Acad. Sci. USA* 60, 1275 (1968)) are suitable for use.

20 The NgR expression vectors prepared as above are introduced into CHO cells by any known method, including, but not limited to the calcium phosphate method (Graham and van der Eb (1973) *Virology* 52, 456-467) and electroporation (Nuemann *et al.*, (1982) *EMBO J.* 1, 841-845).

Transformants carrying the expression vectors are selected based on the above-mentioned selectable markers. Repeated clonal selection of the transformants
25 using the selectable markers allows selection of stable cell lines having high expression of the NgR constructs. Increased MTX concentrations in the selection medium allows gene amplification and greater expression of the desired protein. The CHO cell containing the recombinant NgR can be produced by cultivating the CHO cells containing the NR expression vectors constitutively expressing the NgR constructs.

30 Media used in cultivating CHO cells includes DMEM medium supplemented with about 0.5 to 20% fetal calf serum, DMEM medium and RPMI1640 medium. The

- 30 -

pH of the medium is preferably about 6 to 8. Cultivation is preferably at about 30 to 40°C for about 15 to 72 hours with aeration.

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) NgR protein. Accordingly, the invention further provides methods for producing NgR protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding NgR has been introduced) in a suitable medium such that NgR protein is produced. In another embodiment, the method further comprises isolating NgR from the medium or the host cell.

In situations where the NgR polypeptide will be found primarily intracellularly, intracellular material (including inclusion bodies for Gram-negative bacteria) can be extracted from the host cell using any standard technique known to one of ordinary skill in the art. Such methods would encompass, by way of example and not by way of limitation, lysing the host cells to release the contents of the periplasm/cytoplasm by French press, homogenization and/or sonication followed by centrifugation.

If the NgR polypeptide has formed inclusion bodies in the cytosol, such inclusion bodies may frequently bind to the inner and/or outer cellular membranes. Upon centrifugation, the inclusion bodies will be found primarily in the pellet material. The pellet material can then be treated at pH extremes or with one or more chaotropic agents such as a detergent, guanidine, guanidine derivatives, urea, or urea derivatives in the presence of a reducing agent such as dithiothreitol at alkaline pH or tris-carboxyethyl phosphine at acid pH to release, break apart and solubilize the inclusion bodies. Once solubilized, NgR polypeptide can be analyzed using gel electrophoresis, immunoprecipitation or the like. Various methods of isolating the NgR polypeptide would be apparent to one of ordinary skill in the art, for example, isolation may be accomplished using standard methods such as those set forth below and in Marston *et al* (1990) *Meth. Enzymol.* 182, 264-275 (incorporated by reference herein in its entirety).

If isolated NgR polypeptide is not biologically active following the isolation procedure employed, various methods for "refolding" or converting the polypeptide to its tertiary structure and generating disulfide linkages, can be used to restore biological

- 31 -

activity. Methods known to one of ordinary skill in the art include adjusting the pH of the solubilized polypeptide to a pH usually above 7 and in the presence of a particular concentration of a chaotrope. The selection of chaotrope is very similar to the choices used for inclusion body solubilization but usually at a lower concentration and is not necessarily the same chaotrope as used for the solubilization. It may be required to employ a reducing agent or the reducing agent plus its oxidized form in a specific ratio, to generate a particular redox potential allowing for disulfide shuffling to occur in the formation of the protein's cysteine bridge(s). Some of the commonly used redox couples include cysteine/cystamine, glutathione (GSH)/dithiobis GSH, cupric chloride, dithiothreitol (DTT)/dithiane DTT, 2-mercaptoethanol (bME)/dithio-b(ME). To increase the efficiency of the refolding, it may be necessary to employ a cosolvent, such as glycerol, polyethylene glycol of various molecular weights and arginine.

Transgenic Animals

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which NgR-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous NgR sequences have been introduced into their genome or homologous recombinant animals in which endogenous NgR sequences have been altered. Such animals are useful for studying the function and/or activity of NgR and for identifying and/or evaluating modulators of NgR activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous NgR gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA

- 32 -

molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing NgR-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by
5 microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human NgR DNA sequence of SEQ ID NOs:1 or 3 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homolog of the human NgR gene, such as a mouse NgR gene, can be isolated based on hybridization to the human NgR cDNA (described
10 further above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the NgR transgene to direct expression of NgR protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly
15 animals such as mice, have become conventional in the art and are described, for example, in U.S. Pat. Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan 1986, in MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the
20 NgR transgene in its genome and/or expression of NgR mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding NgR can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which
25 contains at least a portion of a NgR gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the NgR gene. The NgR gene can be a human gene (*e.g.*, SEQ ID NOs:1 or 13), but more preferably, is a non-human homolog of a human NgR gene. For example, a mouse homolog of human NgR gene of SEQ ID NOs:1 or 13 can be used to construct a homologous
30 recombination vector suitable for altering an endogenous NgR gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous

- 33 -

recombination, the endogenous NgR gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous NgR gene is mutated or otherwise altered but still
5 encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous NgR protein). In the homologous recombination vector, the altered portion of the NgR gene is flanked at its 5' and 3' ends by additional nucleic acid of the NgR gene to allow for homologous recombination to occur between the exogenous NgR gene carried by the vector and an
10 endogenous NgR gene in an embryonic stem cell. The additional flanking NgR nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector. See *e.g.*, Thomas *et al.* (1987) *Cell* 51:503 for a description of homologous recombination vectors. The vector is introduced into an
15 embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced NgR gene has homologously recombined with the endogenous NgR gene are selected (see *e.g.*, Li *et al.* (1992) *Cell* 69:915).

The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras. See *e.g.*, Bradley 1987, In:
20 TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A Practical Approach, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined
25 DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Curr. Opin. Biotechnol.* 2:823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

30 In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For

- 34 -

a description of the cre/loxP recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate

5 expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

10 Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmot *et al.* (1997) *Nature* 385:810-813. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal

15 of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

20

Antisense

Also provided by the invention are antisense polynucleotides that recognize and hybridize to NgR polynucleotides. Full-length and fragment antisense polynucleotides are provided. Fragment antisense molecules of the invention include (i) those that

25 specifically recognize and hybridize to NgR RNA (as determined by sequence comparison of DNA encoding NgR to DNA encoding other known molecules). Identification of sequences unique to NgR encoding polynucleotides can be deduced through use of any publicly available sequence database, and/or through use of commercially available sequence comparison programs. After identification of the

30 desired sequences, isolation through restriction digestion or amplification using any of the various polymerase chain reaction techniques well known in the art can be

- 35 -

performed. Antisense polynucleotides are particularly relevant to regulating expression of NgR by those cells expressing NgR mRNA.

Antisense oligonucleotides, or fragments of a nucleotide sequence set forth in SEQ ID NO:1, 3, 13 or sequences complementary or homologous thereto, derived
5 from the nucleotide sequences of the present invention encoding NgR are useful as diagnostic tools for probing gene expression in various tissues. For example, tissue can be probed *in situ* with oligonucleotide probes carrying detectable groups by conventional autoradiography techniques to investigate native expression of this enzyme or pathological conditions relating thereto. In specific aspects, antisense
10 nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire NgR coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a NgR protein of SEQ ID NO:2, 4 or 14 or antisense nucleic acids complementary to a NgR nucleic acid sequence of SEQ ID NOs:1, 3 or
15 13 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding NgR. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the protein coding region of human
20 NgR corresponds to the coding region SEQ ID NO:1, 3 or 13). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding NgR. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

25 Antisense oligonucleotides are preferably directed to regulatory regions of a nucleotide sequence of SEQ ID NO:1, 3, 13 or mRNA corresponding thereto, including, but not limited to, the initiation codon, TATA box, enhancer sequences, and the like. Given the coding strand sequences encoding NgR disclosed herein (*e.g.*, SEQ ID NO:1, 3 or 13), antisense nucleic acids of the invention can be designed according
30 to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of NgR mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding

- 36 -

or noncoding region of NgR mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of NgR mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be
5 constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the
10 antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-
15 carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-
20 mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-
25 diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

30 The antisense nucleic acid molecules of the invention (preferably oligonucleotides of 10 to 20 nucleotides in length) are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA

- 37 -

and/or genomic DNA encoding a NgR protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. Suppression of NgR expression at either the transcriptional or translational level is useful to generate cellular or animal models for diseases/conditions characterized by aberrant NgR expression. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix.

Phosphorothioate and methylphosphonate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. The antisense oligonucleotides may be further modified by adding poly-L-lysine, transferrin polylysine or cholesterol moieties at their 5' end.

An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier *et al.*, (1987) *Nucleic Acids Res.* 15, 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.*, (1987) *Nucleic Acids Res.* 15, 6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.*, (1987) *FEBS Lett.* 215, 327-330).

The NgR sequences taught in the present invention facilitate the design of novel transcription factors for modulating NgR expression in native cells and animals,

- 38 -

and cells transformed or transfected with NgR polynucleotides. For example, the Cys₂-His₂ zinc finger proteins, which bind DNA via their zinc finger domains, have been shown to be amenable to structural changes that lead to the recognition of different target sequences. These artificial zinc finger proteins recognize specific target sites with high affinity and low dissociation constants, and are able to act as gene switches to modulate gene expression. Knowledge of the particular NgR target sequence of the present invention facilitates the engineering of zinc finger proteins specific for the target sequence using known methods such as a combination of structure-based modeling and screening of phage display libraries (Segal *et al.*, (1999) *Proc. Natl. Acad. Sci. USA* 96, 2758-2763; Liu *et al.*, (1997) *Proc. Natl. Acad. Sci. USA* 94, 5525-5530; Greisman *et al.* (1997) *Science* 275, 657-661; Choo *et al.*, (1997) *J. Mol. Biol.* 273, 525-532). Each zinc finger domain usually recognizes three or more base pairs. Since a recognition sequence of 18 base pairs is generally sufficient in length to render it unique in any known genome, a zinc finger protein consisting of 6 tandem repeats of zinc fingers would be expected to ensure specificity for a particular sequence (Segal *et al.*, (1999), above). The artificial zinc finger repeats, designed based on the promoter of NgR sequences, are fused to activation or repression domains to promote or suppress NgR expression (Liu *et al.*, (1997), above). The promoter of NgR may be obtained by standard methods known to one of ordinary skill in the art with the disclosure contained herein and knowledge of the NgR sequence. Alternatively, the zinc finger domains can be fused to the TATA box-binding factor (TBP) with varying lengths of linker region between the zinc finger peptide and the TBP to create either transcriptional activators or repressors (Kim *et al.*, (1997) *Proc. Natl. Acad. Sci. USA* 94, 3616-3620. Such proteins and polynucleotides that encode them, have utility for modulating NgR expression *in vivo* in both native cells, animals and humans; and/or cells transfected with NgR-encoding sequences. The novel transcription factor can be delivered to the target cells by transfecting constructs that express the transcription factor (gene therapy), or by introducing the protein. Engineered zinc finger proteins can also be designed to bind RNA sequences for use in therapeutics as alternatives to antisense or catalytic RNA methods (McColl *et al.*, (1997) *Proc. Natl. Acad. Sci. USA* 96, 9521-9526); Wu *et al.*, (1995) *Proc. Natl. Acad. Sci. USA* 92, 344-348). The present invention contemplates methods of

- 39 -

designing such transcription factors based on the gene sequence of the invention, as well as customized zinc finger proteins, that are useful to modulate NgR expression in cells (native or transformed) whose genetic complement includes these sequences.

5 **Ribozymes and PNA moieties**

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes,
10 described in Haselhoff and Gerlach (1988) *Nature* 334, 585-591) can be used to catalytically cleave NgR mRNA transcripts to thereby inhibit translation of NgR mRNA. A ribozyme having specificity for a NgR-encoding nucleic acid can be designed based upon the nucleotide sequence of a NgR DNA disclosed herein (*i.e.*, SEQ ID NOs:1, 3 or 13). For example, a derivative of a Tetrahymena L-19 IVS RNA
15 can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a NgR-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, NgR mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*,
20 (1993) *Science* 261, 1411-1418.

Alternatively, NgR gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the NgR (*e.g.*, the NgR promoter and/or enhancers) to form triple helical structures that prevent transcription of the NgR gene in target cells. See generally, Helene (1991) *Anticancer Drug Des.* 6:
25 569-584; Helene. *et al.*, (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *BioEssays* 14, 807-815.

In various embodiments, the nucleic acids of NgR can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate
30 backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.*, (1996) *Bioorg. Med. Chem. Lett.* 4, 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in

- 40 -

which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.*, (1996) above; Perry-O'Keefe *et al.*, (1996) *Proc. Natl. Acad. Sci. USA* 93,14670-14675.

PNAs of NgR can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of NgR can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup (1996), above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.*, (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of NgR can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of NgR can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996), above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996), above and Finn *et al.* (1996) *Nucleic Acids Res.* 24, 3357-3363. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl) amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucleic Acids Res.* 17, 973-988). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996), above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment

- 41 -

and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg. Med. Chem. Lett.* 5:1119-1124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see Letsinger *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86, 6553-6556; Lemaitre *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84, 648-652; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (see, *e.g.*, Krol *et al.*, (1988) *Biotechniques* 6, 958-976) or intercalating agents (see, *e.g.*, Zon (1988) *Pharm. Res.* 5, 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

Automated sequencing methods can be used to obtain or verify the nucleotide sequence of NgR. The NgR nucleotide sequences of the present invention are believed to be 100% accurate. However, as is known in the art, nucleotide sequence obtained by automated methods may contain some errors. Nucleotide sequences determined by automation are typically at least about 90%, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of a given nucleic acid molecule. The actual sequence may be more precisely determined using manual sequencing methods, which are well known in the art. An error in a sequence which results in an insertion or deletion of one or more nucleotides may result in a frame shift in translation such that the predicted amino acid sequence will differ from that which would be predicted from the actual nucleotide sequence of the nucleic acid molecule, starting at the point of the mutation.

Polypeptides

The invention also provides purified and isolated mammalian NgR polypeptides encoded by a polynucleotide of the invention. Presently preferred is a human NgR polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:14. Another preferred embodiment is a mouse NgR polypeptide comprising the amino acid sequence of NgR3, as set forth in SEQ ID NO:4.

- 42 -

One aspect of the invention pertains to isolated NgR proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-NgR antibodies. Preferably, fragments of NgR proteins comprise at least one biological activity of NgR. In one embodiment, native NgR proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, NgR proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a NgR protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

10 The invention also embraces polypeptides that have at least 99%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55%, at least 50% or at least 45% identity and/or homology to the preferred polypeptide of the invention. In addition, the invention embraces polypeptides having the consensus sequence shown in SEQ ID NO:6, shown in Table 15 5) excluding the previously characterized NgR ("NgR1"), and polypeptides comprising at least about 90% of the consensus sequence.

The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, 20 dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent 25 identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region.

In one aspect, percent homology is calculated as the percentage of amino acid residues in the smaller of two sequences which align with identical amino acid residue 30 in the sequence being compared, when four gaps in a length of 100 amino acids may be introduced to maximize alignment (Dayhoff, in ATLAS OF PROTEIN SEQUENCE AND

- 43 -

STRUCTURE, Vol. 5, p. 124, National Biochemical Research Foundation, Washington, D.C. (1972), incorporated herein by reference).

A determination of homology or identity is typically made by a computer homology program known in the art. An exemplary program is the Gap program
5 (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison, WI) using the default settings, which uses the algorithm of Smith and Waterman (*Adv. Appl. Math.*, 1981, 2, 482-489, which is incorporated herein by reference in its entirety). Employing the GAP software provided in the GCG program package, (see *Needleman and Wunsch* (1970) *J. Mol.*
10 *Biol.* 48, 443-453) the following settings for nucleic acid sequence comparison may be used: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NOs:1, 3 or 13.
15 BestFit was originally written for Version 1.0 by Paul Haeberli from a careful reading of the papers by Needleman and Wunsch (1970), above, and Smith and Waterman (1981), above. The following Bestfit settings for nucleic acid sequence comparison may be used: GAP creation penalty of 8.0 and GAP extension penalty of 2, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of
20 identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99%, with the CDS (encoding) part of the amino acid sequence shown in SEQ ID NOs:2, 4 or 14.

Alternatively, homology may be determined by hybridization analysis wherein a nucleic acid sequence is hybridized to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent
25 conditions. See e.g. Ausubel, *et al.*, (Eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below.

Polypeptides of the invention may be isolated from natural cell sources or may be chemically synthesized, but are preferably produced by recombinant procedures involving host cells of the invention.

30 An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the NgR protein is derived, or substantially free from

- 44 -

chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of NgR protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of NgR protein having less than about 30% (by dry weight) of non-NgR protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-NgR protein, still more preferably less than about 10% of non-NgR protein, and most preferably less than about 5% non-NgR protein. When the NgR protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of NgR protein in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of NgR protein having less than about 30% (by dry weight) of chemical precursors or non-NgR chemicals, more preferably less than about 20% chemical precursors or non-NgR chemicals, still more preferably less than about 10% chemical precursors or non-NgR chemicals, and most preferably less than about 5% chemical precursors or non-NgR chemicals.

Biologically active portions of a NgR protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the NgR protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2, 4 or 14 that include fewer amino acids than the full length NgR proteins, and exhibit at least one activity of a NgR protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the NgR protein. A biologically active portion of a NgR protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length.

A biologically active portion of a NgR protein of the present invention may contain at least one of the features that is conserved between the NgR proteins (*e.g.*, a conserved cysteine as the N-terminus of the mature protein, four conserved cysteines

- 45 -

in the N-terminus before a leucine-rich region, four conserved cysteines C-terminal with respect to a leucine repeat region, eight leucine-rich repeats, and a hydrophobic C-terminus). An alternative biologically active portion of a NgR protein may contain at least two of the above-identified domains. Another biologically active portion of a
5 NgR protein may contain at least three of the above-identified domains. Yet another biologically active portion of a NgR protein of the present invention may contain at least four of the above-identified domains.

Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one
10 or more of the functional activities of a native NgR protein.

In an embodiment, the NgR protein has an amino acid sequence shown in SEQ ID NO:2, 4 or 14. In other embodiments, the NgR protein is substantially homologous to SEQ ID NO:2, 4 or 14 and retains the functional activity of the protein of SEQ ID NO:2, 4 or 14, yet differs in amino acid sequence due to natural allelic variation or
15 mutagenesis, as described in detail below.

Accordingly, in another embodiment, the NgR protein is a protein that comprises an amino acid sequence at least about 45% homologous to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or SEQ ID NO:14 and retains the functional activity of the NgR proteins of SEQ ID NO:2, 4 or 14.

20 Use of mammalian host cells is expected to provide for such post-translational modifications (*e.g.*, glycosylation, truncation, lipidation and phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. Glycosylated and non-glycosylated forms of NgR polypeptides are embraced by the invention.

25 The invention also embraces variant (or analog) NgR polypeptides. In one example, insertion variants are provided wherein one or more amino acid residues supplement a NgR amino acid sequence. Insertions may be located at either or both termini of the protein, or may be positioned within internal regions of the NgR amino acid sequence. Insertional variants with additional residues at either or both termini
30 can include, for example, fusion proteins and proteins including amino acid tags or labels.

- 46 -

Insertion variants include NgR polypeptides wherein one or more amino acid residues are added to a NgR acid sequence or to a biologically active fragment thereof.

Variant products of the invention also include mature NgR products, *i.e.*, NgR products wherein leader or signal sequences are removed, with additional amino
5 terminal residues. The additional amino terminal residues may be derived from another protein, or may include one or more residues that are not identifiable as being derived from specific proteins. NgR products with an additional methionine residue at position -1 (Met⁻¹-NgR) are contemplated, as are variants with additional methionine and lysine residues at positions -2 and -1 (Met⁻²-Lys⁻¹-NgR). Variants of NgR with
10 additional Met, Met-Lys, Lys residues (or one or more basic residues in general) are particularly useful for enhanced recombinant protein production in bacterial host cells.

Polypeptide Variants

The invention also embraces NgR variants having additional amino acid
15 residues which result from use of specific expression systems.

As used herein, a NgR "chimeric protein" or "fusion protein" comprises a NgR polypeptide operatively linked to a non-NgR polypeptide. A "NgR polypeptide" refers to a polypeptide having an amino acid sequence corresponding to NgR, whereas a "non-NgR polypeptide" refers to a polypeptide having an amino acid sequence
20 corresponding to a protein that is not homologous to the NgR protein, *e.g.*, a protein that is different from the NgR protein and that is derived from the same or a different organism. Within a NgR fusion protein the NgR polypeptide can correspond to all or a portion of a NgR protein. In one embodiment, a NgR fusion protein comprises at least one biologically active portion of a NgR protein. In another embodiment, a NgR
25 fusion protein comprises at least two biologically active portions of a NgR protein. In yet another embodiment, a NgR fusion protein comprises at least three biologically active portions of a NgR protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the NgR polypeptide and the non-NgR polypeptide are fused in-frame to each other. The non-NgR polypeptide can be fused to the N-
30 terminus or C-terminus of the NgR polypeptide.

For example, in one embodiment a NgR fusion protein comprises a NgR domain operably linked to the extracellular domain of a second protein. Such fusion

- 47 -

proteins can be further utilized in screening assays for compounds which modulate NgR activity (such assays are described in detail below).

For example, use of commercially available vectors that express a desired polypeptide as part of a glutathione-S-transferase (GST) fusion product provides the
5 desired polypeptide having an additional glycine residue at position -1 after cleavage of the GST component from the desired polypeptide.

In another embodiment, the fusion protein is a NgR protein containing a heterologous signal sequence at its N-terminus. For example, the native NgR signal sequence (*i.e.*, amino acids 1-30 of SEQ ID NO:2 and amino acids 1-40 of SEQ ID
10 NO:4) can be removed and replaced with a signal sequence from another protein. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion NgR can be increased through use of a heterologous signal sequence.

In yet another embodiment, the fusion protein is a NgR-immunoglobulin fusion protein in which the NgR sequences comprising one or more domains are fused to
15 sequences derived from a member of the immunoglobulin protein family. The NgR-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between NgR ligand and a NgR protein on the surface of a cell, to thereby suppress NgR-mediated signal transduction *in vivo*. NgR-immunoglobulin fusion proteins can be
20 used to affect the bioavailability of a NgR cognate ligand. Inhibition of the NgR ligand/NgR interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the NgR-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-NgR antibodies in a subject, to
25 purify NgR ligands, and in screening assays to identify molecules that inhibit the interaction of NgR with NgR ligand.

A NgR chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional
30 techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining and

- 48 -

enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can
5 subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel *et al.* (Eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A NgR-encoding nucleic acid can be cloned into such an expression vector such that the fusion
10 moiety is linked in-frame to the NgR protein.

Variants resulting from expression in other vector systems are also contemplated.

Insertional variants also include fusion proteins wherein the amino terminus and/or the carboxy terminus of NgR is/are fused to another polypeptide.

15 In another aspect, the invention provides deletion variants wherein one or more amino acid residues in a NgR polypeptide are removed. Deletions can be effected at one or both termini of the NgR polypeptide, or with removal of one or more non-terminal amino acid residues of NgR. Deletion variants, therefore, include all fragments of a NgR polypeptide.

20 The invention also embraces polypeptide fragments of the sequence set forth in SEQ ID NO:2, 4 or 14 wherein the fragments maintain biological (*e.g.*, ligand binding and/or intracellular signaling) immunological properties of a NgR polypeptide. Fragments comprising at least 4, 5, 10, 15, 20, 25, 30, 35, or 40 consecutive amino acids of SEQ ID NO:2, 4 or 14 are contemplated by the invention. Preferred
25 polypeptide fragments display antigenic properties unique to, or specific for, human NgR and its allelic and species homologs. Fragments of the invention having the desired biological and immunological properties can be prepared by any of the methods well known and routinely practiced in the art.

In still another aspect, the invention provides substitution variants of NgR
30 polypeptides. Substitution variants include those polypeptides wherein one or more amino acid residues of a NgR polypeptide are removed and replaced with alternative residues. In one aspect, the substitutions are conservative in nature; however, the

invention embraces substitutions that are also non-conservative. Conservative substitutions for this purpose may be defined as set out in Tables 2, 3, or 4 below.

Table 1.

5	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₁	G, R, M	
	X ₂	A, D, C	
	X ₃	V, T	
10	X ₄	N, P, S	
	X ₅	E, A, S	
	X ₆	nothing, K	nothing
	X ₇	V, M, P	
	X ₈	T, V	V
15	X ₉	Q, P	Q
	X ₁₀	Q, A	Q
	X ₁₁	Q, H, N	
	X ₁₂	G, N	N
	X ₁₃	L, F	F
20	X ₁₄	Q, A, S	
	X ₁₅	A, S	
	X ₁₆	V, I	
	X ₁₇	V, T, E, L	
	X ₁₈	S, G	
25	X ₁₉	L, I	
	X ₂₀	A, E, V, P	
	X ₂₁	A, S, D	
	X ₂₂	S, T	
	X ₂₃	Q, E	

- 50 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₂₄	IVL	
	X ₂₅	Q,H	Q
	X ₂₆	N,G	N
	X ₂₇	R,L	
5	X ₂₈	T,G,R,S	
	X ₂₉	F,L,T,H	
	X ₃₀	L,V	L
	X ₃₁	Q,R,P	
	X ₃₂	Q,P,A	P
10	X ₃₃	G,A	G
	X ₃₄	H,T,S	
	X ₃₅	S,G,R	
	X ₃₆	P,S,A	
	X ₃₇	C, nothing	nothing
15	X ₃₈	R, nothing	nothing
	X ₃₉	A,N	
	X ₄₀	M,L	
	X ₄₁	V,L,T	
	X ₄₂	T, I	T
20	X ₄₃	L,I	
	X ₄₄	Y,F,H	
	X ₄₅	N,V	N
	X ₄₆	I,L	
	X ₄₇	T,S,A	
25	X ₄₈	F,Y,T,R	
	X ₄₉	A,H,Y,D	

- 51 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₅₀	P,A	P
	X ₅₁	N,S,G,A	
	X ₅₂	T,A	T
	X ₅₃	E,R,T	
5	X ₅₄	G,H	
	X ₅₅	F,L	
	X ₅₆	V,Q,H	
	X ₅₇	H,A,L	
	X ₅₈	E,Q	E
10	X ₅₉	G,S	G
	X ₆₀	R,A	R
	X ₆₁	Q,H	Q
	X ₆₂	R, H	H
	X ₆₃	T,S	
15	X ₆₄	L,V	L
	X ₆₅	A,E,D	
	X ₆₆	E,D,A	
	X ₆₇	Q,H	Q
	X ₆₈	V,E,G	
20	X ₆₉	K,R	
	X ₇₀	H,Q	
	X ₇₁	A,S,T	
	X ₇₂	Y,H	
	X ₇₃	Y,D	Y
25	X ₇₄	K,R	
	X ₇₅	G,Q	

- 52 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₇₆	S,Q	S
	X ₇₇	A,S,E	
	X ₇₈	P,G	P
	X ₇₉	A,G,P	
5	X ₈₀	G,N	
	X ₈₁	I,V,L	
	X ₈₂	G,R	
	X ₈₃	H,V,A	
	X ₈₄	S,A	S
10	X ₈₅	D,E	
	X ₈₆	H,S,A	
	X ₈₇	I,L	
	X ₈₈	E,L,Q	
	X ₈₉	Y,H,A	
15	X ₉₀	Q,P	Q
	X ₉₁	D, N	
	X ₉₂	I, L, T	
	X ₉₃	V,A,R	
	X ₉₄	V,A,G	
20	X ₉₅	S,T	S
	X ₉₆	K,R	
	X ₉₇	L,I	L
	X ₉₈	W,R,S	
	X ₉₉	S,L	
25	X ₁₀₀	L,V	L
	X ₁₀₁	G,T,P	

- 53 -

	X_{an}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₁₀₂	Q,P,E	
	X ₁₀₃	G,H,R	
	X ₁₀₄	I,T,V,A	
	X ₁₀₅	V,G,H	
5	X ₁₀₆	N,S	
	X ₁₀₇	E,G,Q	
	X ₁₀₈	Q,R	
	X ₁₀₉	L,V	
	X ₁₁₀	Q,A	
10	X ₁₁₁	W,G,H	
	X ₁₁₂	H,R,P	
	X ₁₁₃	K,A,H	
	X ₁₁₄	H,R	
	X ₁₁₅	D,G	
15	X ₁₁₆	H,R,S,G	
	X ₁₁₇	T,M	
	X ₁₁₈	T,I	
	X ₁₁₉	F,Y	
	X ₁₂₀	N,A	
20	X ₁₂₁	S,N	
	X ₁₂₂	T,A,S	
	X ₁₂₃	E,S,A	
	X ₁₂₄	Q,P	
	X ₁₂₅	G,T	
25	X ₁₂₆	D,E	D
	X ₁₂₇	C,A	

- 54 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₁₂₈	P,D	
	X ₁₂₉	V,G,P,R	
	X ₁₃₀	A,S	
	X ₁₃₁	E,Q	Q
5	X ₁₃₂	F,Y	F
	X ₁₃₃	G,A,D	
	X ₁₃₄	A,P	
	X ₁₃₅	D,A,V	
	X ₁₃₆	G,D	
10	X ₁₃₇	A,E	
	X ₁₃₈	S,P	
	X ₁₃₉	E,A	
	X ₁₄₀	L,F	
	X ₁₄₁	R,Q	
15	X ₁₄₂	R,K	R
	X ₁₄₃	R,K	R
	X ₁₄₄	F,A	
	X ₁₄₅	G,V	
	X ₁₄₆	A,D,E	
20	X ₁₄₇	T,P	
	X ₁₄₈	A,V,S	
	X ₁₄₉	T,S,L	
	X ₁₅₀	E,G,P,Q	
	X ₁₅₁	L,E,R	
25	X ₁₅₂	R,L	R
	X ₁₅₃	G, D	

- 55 -

X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
X ₁₅₄	Q,H,A	
X ₁₅₅	Q,R	
X ₁₅₆	K,R	
X ₁₅₇	L,A,R	
X ₁₅₈	R,A	R
X ₁₅₉	V,A,E	
X ₁₆₀	E,A,N	
X ₁₆₁	F,L	F
X ₁₆₂	R,Q	
X ₁₆₃	N,A,G	

Variant polypeptides include those wherein conservative substitutions have been introduced by modification of polynucleotides encoding polypeptides of the invention. Amino acids can be classified according to physical properties and contribution to secondary and tertiary protein structure. A conservative substitution is recognized in the art as a substitution of one amino acid for another amino acid that has similar properties. Exemplary conservative substitutions are set out in Table 2 (from WO 97/09433, page 10, published March 13, 1997 (PCT/GB96/02197, filed 9/6/96), immediately below.

Table 2

Conservative Substitutions I

SIDE CHAIN CHARACTERISTIC	AMINO ACID
Aliphatic	G A P
Non-polar	I L V

- 56 -

Polar - uncharged

C S T M
N Q

Polar - charged

D E
K R

Aromatic

H F W Y

Other

N Q D E

5 Alternatively, conservative amino acids can be grouped as described in
Lehninger, [BIOCHEMISTRY, Second Edition; Worth Publishers, Inc. NY, NY
(1975), pp.71-77] as set out in Table 3, immediately below.

Table 3

Conservative Substitutions II

10

SIDE CHAIN**CHARACTERISTIC****AMINO ACID**

Non-polar (hydrophobic)

A. Aliphatic:

A L I V P

B. Aromatic:

F W

15

C. Sulfur-containing:

M

D. Boderline:

G

Uncharged-polar

A. Hydroxyl:

S T Y

B. Amides:

N Q

C. Sylfhydryl:

C

20

D. Boderline:

G

Positively Charged (Basic):

K R H

Negatively Charged (Acidic):

D E

As still another alternative, exemplary conservative substitutions are set out in
25 Table 4, below.

Table 4

Conservative Substitutions III**Original
Residue****Exemplary Substitution**

30

Ala (A)

Val, Leu, Ile

Arg (R)

Lys, Gln, Asn

Asn (N)

Gln, His, Lys, Arg

Asp (D)

Glu

Cys (C)

Ser

Gln (Q)

Asn

35

Glu (E)

Asp

- 57 -

	His (H)	Asn, Gln, Lys, Arg
	Ile (I)	Leu, Val, Met, Ala, Phe,
	Leu (L)	Ile, Val, Met, Ala, Phe
	Lys (K)	Arg, Gln, Asn
5	Met (M)	Leu, Phe, Ile
	Phe (F)	Leu, Val, Ile, Ala
	Pro (P)	Gly
	Ser (S)	Thr
	Thr (T)	Ser
10	Trp (W)	Tyr
	Tyr (Y)	Trp, Phe, Thr, Ser.
	Val (V)	Ile, Leu, Met, Phe, Ala

In addition, amino acid residues that are conserved among family members of the NgR proteins of the present invention, as indicated by the alignment presented herein, are also predicted to be particularly unamenable to alteration. For example, NgR proteins of the present invention can contain at least one domain that is a typically conserved region in NgRs. Examples of these conserved domains include, *e.g.*, leucine-rich repeat domain. Amino acid residues that are not conserved or are only semi-conserved among members of the NgR proteins may be readily amenable to alteration.

Full-length NgRs have an LRR region characterized by the amino acid consensus sequence shown in SEQ ID NO: 19. At least some full-length NgRs also include a CT signaling (CTS) domain and a GPI domain.

The NgR domain designations used herein are defined as follows:

Domain	hNgR1 SEQ ID: 5	mNgR1 SEQ ID NO:17	hNgR2 SEQ ID: 2	hNgR3 SEQ ID: 14	mNgR3 SEQ ID: 4
Signal Seq.	1-26	1-26	1-30	—	1-40
LRRNT	27-56	27-56	31-59	—	41-69
LRR1	57-81	57-81	60-82	5-27	70-92
LRR2	82-105	82-105	83-106	28-51	93-106
LRR3	106-130	106-130	107-131	52-76	106-141
LRR4	131-154	131-154	132-155	77-100	142-165

- 58 -

	LRR5	155-178	155-178	156-179	101-124	166-189
	LRR6	179-202	179-202	180-203	125-148	190-213
	LRR7	203-226	203-226	204-227	149-172	214-237
	LRR8	227-250	227-250	228-251	173-196	238-261
5	LRRCT	260-309	260-309	261-310	206-255	271-320
	CTS (CT Signaling)	310-445	310-445	311-395	256-396	321-438
	GPI	446-473	456-473	396-420	370-392	439-462

10

In some embodiments of the invention, the above domains are modified. Modification can be in a manner that preserves domain functionality. Modification can include addition, deletion or substitution of certain amino acids. Exemplary modifications include conservative amino acid substitutions. Preferably such substitutions number 20 or fewer per 100 residues. More preferably, such substitutions number 10 or fewer per 100 residues. Further exemplary modifications include addition of flanking sequences of up to five amino acids at the N terminus and/or C terminus of one or more of the domains.

15

In some embodiments, the isolated nucleic acid molecule encodes a polypeptide at least about 70%, 80%, 90%, 95%, 98%, and most preferably at least about 99% homologous to SEQ ID NO:2, 4 or 14.

20

Mutations can be introduced into SEQ ID NOS:1, 3 or 13 by standard techniques, e.g., site-directed mutagenesis and PCR-mediated mutagenesis. Conservative amino acid substitutions can be made at one or more amino acid residues predicted to be non-essential. Alternatively, mutations can be introduced randomly along a NgR coding sequence. This can be accomplished, e.g., by saturation mutagenesis. The resulting mutants can be screened for NgR biological activity. Biological activities of NgR may include but are not limited to: (1) protein:protein interactions, e.g., with other NgRs or other cell-surface proteins involved in Nogo-related signaling; (2) complex formation with a NgR ligand; (3) binding to an anti-NgR antibody.

25

30

- 59 -

It should be understood that the definition of polypeptides of the invention is intended to include polypeptides bearing modifications other than insertion, deletion, or substitution of amino acid residues. By way of example, the modifications may be covalent in nature, and include for example, chemical bonding with polymers, lipids, other organic and inorganic moieties. Such derivatives may be prepared to increase circulating half-life of a polypeptide, or may be designed to improve the targeting capacity of the polypeptide for desired cells, tissues or organs. Similarly, the invention further embraces NgR polypeptides that have been covalently modified to include one or more water-soluble polymer attachments such as polyethylene glycol, polyoxyethylene glycol or polypropylene glycol. Variants that display ligand binding properties of native NgR and are expressed at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant NgR activity.

Chemically modified NgR polypeptide compositions in which the NgR polypeptide is linked to a polymer are included within the scope of the present invention. The polymer may be water soluble to prevent precipitation of the protein in an aqueous environment, such as a physiological environment. Suitable water-soluble polymers may be selected from the group consisting of, for example, polyethylene glycol (PEG), monomethoxypolyethylene glycol, dextran, cellulose, or other carbohydrate based polymers, poly-(N-vinyl pyrrolidone) polyethylene glycol, polypropylene glycol homopolymers, a polypropylene oxide/ethylene oxide copolymer polyoxyethylated polyols (*e.g.* glycerol) and polyvinyl alcohol. The selected polymer is usually modified to have a single reactive group, such as an active ester for acylation or an aldehyde for alkylation, so that the degree of polymerization may be controlled. Polymers may be of any molecular weight, and may be branched or unbranched, and mixtures of such polymers may also be used. When the chemically modified NgR polymer is destined for therapeutic use, pharmaceutically acceptable polymers will be selected for use.

When the polymer is to be modified by an acylation reaction, the polymer should have a single reactive ester group. Alternatively, if the polymer is to be modified by reductive alkylation, the polymer should have a single reactive aldehyde

- 60 -

group. A preferred reactive aldehyde is polyethylene glycol propionaldehyde, which is water stable, or mono Cl-CIO alkoxy or aryloxy derivatives thereof (see U.S. Patent No. 5,252,714, incorporated by reference herein in its entirety).

Pegylation of NgR polypeptides may be carried out by any of the pegylation reactions known in the art, as described, for example, in the following references: *Focus on Growth Factors* 3, 4-10 (1992); EP 0 154 316; and EP 0 401 384 (each of which is incorporated by reference herein in its entirety). Preferably, the pegylation is carried out via an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer). A preferred water-soluble polymer for pegylation of polypeptides such as NgR is polyethylene glycol (PEG). As used herein, "polyethylene glycol" is meant to encompass any of the forms of PEG that have been used to derivatize other proteins, such as mono (Cl-CIO) alkoxy- or aryloxy-polyethylene glycol.

Chemical derivatization of NgR polypeptides may be performed under any suitable conditions used to react a biologically active substance with an activated polymer molecule. Methods for preparing pegylated NgR polypeptides will generally comprise the steps of (a) reacting the polypeptide with polyethylene glycol, such as a reactive ester or aldehyde derivative of PEG, under conditions whereby NgR polypeptide becomes attached to one or more PEG groups, and (b) obtaining the reaction products. It will be apparent to one of ordinary skill in the art to select the optimal reaction conditions or the acylation reactions based on known parameters and the desired result.

Pegylated and other polymer:NgR polypeptides may generally be used to treat conditions that may be alleviated or modulated by administration of the NgR polypeptides described herein. However, the chemically-derivatized polymer:NgR polypeptide molecules disclosed herein may have additional activities, enhanced or reduced biological activity, or other characteristics, such as increased or decreased half-life, as compared to the nonderivatized molecules. The NgR polypeptides, fragments thereof, variants and derivatives, may be employed alone, together, or in combination with other pharmaceutical compositions. The cytokines, growth factors, antibiotics, antiinflammatories and/or chemotherapeutic agents as is appropriate for the indication being treated.

- 61 -

The present invention provides compositions comprising purified polypeptides of the invention. Preferred compositions comprise, in addition to the polypeptide of the invention, a pharmaceutically acceptable (*i.e.*, sterile and non-toxic) liquid, semisolid, or solid diluent that serves as a pharmaceutical vehicle, excipient or medium.

5 Any diluent known in the art may be used. Exemplary diluents include, but are not limited to, water, saline solutions, polyoxyethylene sorbitan monolaurate, magnesium stearate, methyl- and propylhydroxybenzoate, talc, alginates, starches, lactose, sucrose, dextrose, sorbitol, mannitol, glycerol, calcium phosphate, mineral oil and cocoa butter.

Variants that display ligand binding properties of native NgR and are expressed
10 at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in assays of the invention and in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant NgR activity.

With the knowledge of the nucleotide sequence information disclosed in the
15 present invention, one skilled in the art can identify and obtain nucleotide sequences which encode NgR from different sources (*i.e.*, different tissues or different organisms) through a variety of means well known to the skilled artisan and as disclosed by, for example, Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, Second Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), which is
20 incorporated herein by reference in its entirety.

For example, DNA that encodes NgR may be obtained by screening of mRNA, cDNA, or genomic DNA with oligonucleotide probes generated from the NgR gene sequence information provided herein. Probes may be labeled with a detectable group, such as a fluorescent group, a radioactive atom or a chemiluminescent group in
25 accordance with procedures known to the skilled artisan and used in conventional hybridization assays, as described by, for example, Sambrook *et al.* (1989) above.

A nucleic acid molecule comprising any of the NgR nucleotide sequences described above can alternatively be synthesized by use of the polymerase chain reaction (PCR) procedure, with the PCR oligonucleotide primers produced from the
30 nucleotide sequences provided herein. See U.S. Patent Nos. 4,683,195 to Mullis *et al.* and 4,683,202 to Mullis. The PCR reaction provides a method for selectively increasing the concentration of a particular nucleic acid sequence even when that

- 62 -

sequence has not been previously purified and is present only in a single copy in a particular sample. The method can be used to amplify either single- or double-stranded DNA. The essence of the method involves the use of two oligonucleotide probes to serve as primers for the template-dependent, polymerase-mediated replication of a desired nucleic acid molecule.

A wide variety of alternative cloning and *in vitro* amplification methodologies are well known to those skilled in the art. Examples of these techniques are found in, for example, Berger *et al.*, *Guide to Molecular Cloning Techniques*, METHODS IN ENZYMOLOGY 152 Academic Press, San Diego, CA, which is incorporated herein by reference in its entirety.

The nucleic acid molecules of the present invention, and fragments derived therefrom, are useful for screening for restriction fragment length polymorphism (RFLP) associated with certain disorders, as well as for genetic mapping.

15 Antibodies

Also comprehended by the present invention are antibodies (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR sequences which specifically recognize a polypeptide of the invention) specific for NgR or fragments thereof. Preferred antibodies of the invention are human antibodies which are produced and identified according to methods described in WO93/11236, published June 20, 1993, which is incorporated herein by reference in its entirety. Antibody fragments, including Fab, Fab', F(ab')₂, and F_v, are also provided by the invention. The term "specific for," when used to describe antibodies of the invention, indicates that the variable regions of the antibodies of the invention recognize and bind NgR polypeptides exclusively (*i.e.*, are able to distinguish NgR polypeptides from other known NgR polypeptides by virtue of measurable differences in binding affinity, despite the possible existence of localized sequence identity, homology, or similarity between NgR and such polypeptides).

The antigenic peptide of NgR comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2, 4 or 14 and encompasses an epitope of

- 63 -

NgR such that an antibody raised against the peptide forms a specific immune complex with NgR. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues. Preferred epitopes
5 encompassed by the antigenic peptide are regions of NgR that are located on the surface of the protein, *e.g.*, hydrophilic regions.

It will be understood that specific antibodies may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and,
10 in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow *et al.* in ANTIBODIES: A LABORATORY MANUAL, Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments
15 of the NgR polypeptides of the invention are also contemplated, provided that the antibodies are specific for NgR polypeptides. Antibodies of the invention can be produced using any method well known and routinely practiced in the art.

For the production of polyclonal antibodies, various suitable host animals (*e.g.*, rabbit, goat, mouse or other mammal) may be immunized by injection with the native
20 protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly expressed NgR protein or a chemically synthesized NgR polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (*e.g.*,
25 aluminum hydroxide), surface active substances (*e.g.*, lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as *Bacille Calmette-Guerin* and *Corynebacterium parvum* or similar immunostimulatory agents. If desired, the antibody molecules directed against NgR can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such
30 as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition," as used herein, refers to a population of antibody molecules that contain only one species

- 64 -

of an antigen binding site capable of immunoreacting with a particular epitope of NgR. A monoclonal antibody composition thus typically displays a single binding affinity for a particular NgR protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular NgR protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see Kohler and Milstein (1975) *Nature* 256, 495-497); the trioma technique; the human B-cell hybridoma technique (see Kozbor *et al.*, (1983) *Immunol. Today* 4, 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole *et al.*, (1985) in MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote *et al.*, (1983) *Proc. Natl. Acad. Sci. USA* 80, 2026-2030) or by transforming human B-cells with Epstein Barr Virus *in vitro* (see Cole *et al.*, (1985), above).

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a NgR protein (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of Fab expression libraries (see *e.g.*, Huse *et al.*, (1989) *Science* 246, 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a NgR protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See *e.g.*, U.S. Patent No. 5,225,539. In one method, the non-human CDRs are inserted into a human antibody or consensus antibody framework sequence. Further changes can then be introduced into the antibody framework to modulate affinity or immunogenicity. Antibody fragments that contain the idiotypes to a NgR protein may be produced by techniques known in the art including, but not limited to: (i) an F(ab')₂ fragment produced by pepsin digestion of an antibody molecule; (ii) an Fab fragment generated by reducing the disulfide bridges of an F(ab')₂ fragment; (iii) an Fab fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

- 65 -

Additionally, recombinant anti-NgR antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Pat. No. 4,816,567; European Patent Application No. 125,023; Better *et al.*, (1988) *Science* 240, 1041-1043; Liu *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84, 3439-3443; Liu *et al.*, (1987) *J. Immunol.* 139, 3521-3526; Sun *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84, 214-218; Nishimura *et al.*, (1987) *Cancer Res.* 47, 999-1005; Wood *et al.*, (1985) *Nature* 314, 446-449; Shaw *et al.*, (1988) *J. Natl. Cancer Inst.* 80, 1553-1559; Morrison (1985) *Science* 229, 1202-1207; Oi *et al.*, (1986) *BioTechniques* 4, 214; U.S. Patent. No. 5,225,539; Jones *et al.*, (1986) *Nature* 321, 552-525; Verhoeyan *et al.*, (1988) *Science* 239, 1534; and Beidler *et al.*, (1988) *J. Immunol.* 141, 4053-4060.

In a preferred embodiment of the invention a portion of a NgR is joined to an Fc portion of an antibody to form a NgR/Fc fusion protein. Preferably, the Ig fusion protein is soluble. The NgR/Fc fusion protein may be formed by recombinant techniques as described above. In one embodiment, a portion of a NgR including the entire amino acid sequence of NgR except the C-terminal hydrophobic region is fused to an Fc portion of an antibody. In preferred embodiments, the NgR is a human NgR and the Fc is also human. More preferably, the human Fc portion is derived from an IgG antibody. In other embodiments, the N-terminal signal sequence is omitted. Such antibodies are useful in binding Nogo to prevent Nogo signaling through the NgR.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of a NgR protein is facilitated by generation of hybridomas that bind to the fragment of a NgR protein possessing such a domain. Antibodies that are specific for one or more

- 66 -

domains within a NgR protein, *e.g.*, domains spanning the above-identified conserved regions of NgRs, or derivatives, fragments analogs or homologs thereof, are also provided herein.

Anti-NgR antibodies may be used in methods known within the art relating to the localization and/or quantitation of a NgR protein (*e.g.*, for use in measuring levels of the NgR protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for NgR proteins, or derivatives, fragments analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds [hereinafter "Therapeutics"].

An anti-NgR antibody (*e.g.*, monoclonal antibody) can be used to isolate NgR by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-NgR antibody can facilitate the purification of natural NgR from cells and of recombinantly produced NgR expressed in host cells. Moreover, an anti-NgR antibody can be used to detect NgR protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the NgR protein. Anti-NgR antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

Another aspect of the present invention is directed to methods of inducing an immune response in a mammal against a polypeptide of the invention by administering

- 67 -

to the mammal an amount of the polypeptide sufficient to induce an immune response. The amount will be dependent on the animal species, size of the animal, and the like but can be determined by those skilled in the art.

Another aspect of the invention is directed to anti-idiotypic antibodies and
5 anti-anti-idiotypic antibodies. An anti-idiotypic antibody is an antibody that recognizes determinants of another antibody (a target antibody). Generally, the anti-idiotypic antibody recognizes determinants of the antigen-binding site of the target antibody. Typically, the target antibody is a monoclonal antibody. An anti-idiotypic antibody is generally prepared by immunizing an animal (particularly, mice) of the same species
10 and genetic type as the source of the target monoclonal antibody, with the target monoclonal antibody. The immunized animal mounts an immune response to the idiotypic determinants of the target monoclonal antibody and produces antibodies against the idiotypic determinants of the target monoclonal antibody. Antibody-producing cells, such as splenic cells, of the immunized animal may be used
15 to generate anti-idiotypic monoclonal antibodies. Furthermore, an anti-idiotypic antibody may also be used to immunize animals to produce anti-anti-idiotypic antibodies. These immunized animals may be used to generate anti-anti-idiotypic monoclonal antibodies using standard techniques. The anti-anti-idiotypic antibodies may bind to the same epitope as the original, target monoclonal antibody used to
20 prepare the anti-idiotypic antibody. The anti-anti-idiotypic antibodies represent other monoclonal antibodies with the same antigen specificity as the original target monoclonal antibody.

If the binding of the anti-idiotypic antibody with the target antibody is inhibited by the relevant antigen of the target antibody, and if the anti-idiotypic antibody induces
25 an antibody response with the same specificity as the target antibody, it mimics the antigen of the target antibody. Such an anti-idiotypic antibody is an "internal image anti-idiotype" and is capable of inducing an antibody response as if it were the original antigen. (Bona and Kohler (1984) ANTI-IDIOTYPIC ANTIBODIES AND INTERNAL IMAGE, IN MONOCLONAL AND ANTI-IDIOTYPIC ANTIBODIES: PROBES FOR RECEPTOR STRUCTURE
30 AND FUNCTION, Venter J.C. *et al.* (Eds), Alan R. Liss, New York, NY, pp 141-149, 1984). Vaccines incorporating internal image anti-idiotype antibodies have been shown to induce protective responses against viruses, bacteria, and parasites (Kennedy

- 68 -

et al., (1986) 232, 220-223; 1047; McNamara *et al.*, (1985) *Science* 226, 1325-1326). Internal image anti-idiotypic antibodies have also been shown to induce immunity to tumor related antigens (Raychauhuri *et al.*, (1986) *J. Immunol.* 137, 1743-1749; Raychauhuri *et al.*, (1987) *J. Immunol.* 139, 3902-3910; Bhattacharya-Chatterjee *et al.*, (1987) *J. Immunol.* 139, 1354-1360; Bhattacharya-Chatterjee *et al.*, (1988) *J. Immunol.* 141, 1398-1403; Herlyn. *et al.* (1989) *Intern. Rev. Immunol.* 4, 347-357; Chen *et al.* (1990) *Cell Imm. Immunother. Cancer* 351-359; Herlyn *et al.*, (1991) *in vivo* 5, 615-624; Furuya *et al.* (1992) *AntiCancer Res.* 12, 27-32; Mittelman, A. *et al.* (1992) *Proc. Natl. Acad. Sci., USA* 89, 466-470; Durrant. *et al.*, (1994) *Cancer Res.* 54, 4837-4840; Mittelman. *et al.* (1994) *Cancer Res.* 54, 415-421; Schmitt. *et al.* (1994) *Hybridoma* 13, 389-396; Chakrobarty. *et al.* (1995) *J. Immunother.* 18, 95-103; Chakrobarty. *et al.* (1995) *Cancer Res.* 55, 1525-1530; Foon, K. A. *et al.* (1995) *Clin. Cancer Res.* 1, 1205-1294; Herlyn *et al.* (1995) *Hybridoma* 14, 159-166; Sclebusch *et al.* (1995) *Hybridoma* 14, 167-174; Herlyn. *et al.* (1996) *Cancer Immunol Immunother.* 43, 65-76).

Anti-idiotypic antibodies for NgR may be prepared, for example, by immunizing an animal, such as a mouse, with a immunogenic amount of a composition comprising NgR2 (SEQ ID NO:2), NgR3 (SEQ ID NOs:4 or 14), or immunogenic portion thereof, containing at least one antigenic epitope of NgR. The composition may also contain a suitable adjuvant, and any carrier necessary to provide immunogenicity. Monoclonal antibodies recognizing NgR may be prepared from the cells of the immunized animal as described above. A monoclonal antibody recognizing an epitope of NgR is then selected and used to prepare a composition comprising an immunogenic amount of the anti-NgR monoclonal antibody. Typically, a 25 to 200 µg dose of purified anti-NgR monoclonal would be sufficient in a suitable adjuvant.

Animals may be immunized 2-6 times at 14 to 30 day intervals between doses. Typically, animals are immunized by any suitable route of administration, such as intraperitoneal, subcutaneous, intravenous or a combination of these. Anti-idiotypic antibody production may be monitored during the immunization period using standard immunoassay methods. Animals with suitable titers of antibodies reactive with the target monoclonal antibodies may be reimmunized with the monoclonal antibody used as the immunogen three days before harvesting the antibody producing cells.

- 69 -

Preferably, spleen cells are used, although other antibody producing cells may be selected. Antibody-producing cells are harvested and fused with myeloma cells to produce *Hybridomas*, as described above, and suitable anti-idiotypic antibody-producing cells are selected.

- 5 Anti-anti-idiotypic antibodies are produced by another round of immunization and *Hybridoma* production by using the anti-idiotypic monoclonal antibody as the immunogen.

- Antibodies of the invention are useful for, *e.g.*, therapeutic purposes (by modulating activity of NgR), diagnostic purposes to detect or quantitate NgR, and
10 purification of NgR. Therefore, kits comprising an antibody of the invention for any of the purposes described herein are also comprehended.

Kits

- The present invention is also directed to kits, including pharmaceutical kits.
- 15 The kits can comprise any of the nucleic acid molecules described above, any of the polypeptides described above, or any antibody which binds to a polypeptide of the invention as described above, as well appropriate controls, such as positive and/or negative controls. The kit preferably comprises additional components, such as, for example, instructions, solid support, reagents helpful for quantification, and the like.
- 20 For example, the kit can comprise: a labeled compound or agent capable of detecting NgR protein or mRNA in a biological sample; means for determining the amount of NgR in the sample; and means for comparing the amount of NgR in the sample with a standard. The compound or agent can be packaged in a suitable container.

25 Screening Assays

- The DNA and amino acid sequence information provided by the present invention also makes possible identification of binding partner compounds with which a NgR polypeptide or polynucleotide will interact. Methods to identify binding partner compounds include solution assays, *in vitro* assays wherein NgR polypeptides are
30 immobilized and cell-based assays. Identification of binding partner compounds of NgR polypeptides provides candidates for therapeutic or prophylactic intervention in pathologies associated with NgR normal and aberrant biological activity.

- 70 -

The invention also provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules (*e.g.*, molecules of less than 1,000 Daltons) or other drugs) that bind to NgR proteins or have a stimulatory or inhibitory effect on, for example, NgR expression or NgR activity.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a NgR protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12, 145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.*, (1993) *Proc. Natl. Acad. Sci. USA* 90, 6909; Erb *et al.*, (1994) *Proc. Natl. Acad. Sci. USA* 91, 11422; Zuckermann *et al.* (1994) *J. Med. Chem.* 37, 2678; Cho *et al.*, (1993) *Science* 261, 1303; Carrell *et al.*, (1994) *Angew Chem. Int. Ed. Engl.* 33, 2059; Carell *et al.*, (1994) *Angew Chem. Int. Ed. Engl.* 33, 2061; and Gallop *et al.*, (1994) *J. Med. Chem.* 37, 1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *BioTechniques* 13, 412-421), or on beads (Lam (1991) *Nature* 354, 82-84), on chips (Fodor (1993) *Nature* 364, 555-556), bacteria (Ladner, U.S. Patent No. 5,223,409), spores (Ladner, above), plasmids (Cull *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89, 1865-1869) or on phage (Scott and Smith (1990) *Science* 249, 386-390; Devlin (1990) *Science* 249, 404-406; Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87, 6378-6382; Felici (1991) *J. Mol. Biol.* 222, 301-310; Ladner, above).

1. Cell-based Assays

The invention also provides cell-based assays to identify binding partner compounds of a NgR polypeptide. In one embodiment, the invention provides a method comprising the steps of contacting a NgR polypeptide expressed on the surface of a cell with a candidate binding partner compound and detecting binding of the candidate binding partner compound to the NgR polypeptide. In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of NgR protein, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the NgR protein or biologically active portion thereof.

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of NgR protein, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to a NgR protein determined. The cell, for example, can be of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the NgR protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the NgR protein or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of NgR protein or a biologically active portion thereof, on the cell surface with a known compound which binds NgR to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NgR protein, wherein determining the ability of the test compound to interact with a NgR protein comprises determining the ability of the test compound to preferentially bind to NgR or a biologically active portion thereof as compared to the known compound.

- 72 -

Determining the ability of the test compound to modulate the activity of NgR or a biologically active portion thereof can be accomplished, for example, by determining the ability of the NgR protein to bind to or interact with a NgR target molecule. As used herein, a "target molecule" is a molecule with which a NgR protein
5 binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a NgR protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. A NgR target molecule can be a non-NgR molecule or a NgR protein or polypeptide of the present invention. In one embodiment, a NgR
10 target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.*, a signal generated by binding of a compound to a membrane-bound NgR molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling
15 molecules with NgR. In a preferred embodiment, the detection comprises detecting a calcium flux or other physiological event in the cell caused by the binding of the molecule.

Specific binding molecules, including natural ligands and synthetic compounds, can be identified or developed using isolated or recombinant NgR products, NgR
20 variants, or preferably, cells expressing such products. Binding partners are useful for purifying NgR products and detection or quantification of NgR products in fluid and tissue samples using known immunological procedures. Binding molecules are also manifestly useful in modulating (*i.e.*, blocking, inhibiting or stimulating) biological activities of NgR, especially those activities involved in signal transduction.

25

2. Cell-free Assays

(a) Direct binding:

The invention includes several assay systems for identifying NgR binding partners. In solution assays, methods of the invention comprise the steps of (a)
30 contacting a NgR polypeptide with one or more candidate binding partner compounds and (b) identifying the compounds that bind to the NgR polypeptide. Identification of the compounds that bind the NgR polypeptide can be achieved by isolating the NgR

- 73 -

polypeptide/binding partner complex and separating the binding partner compound from the NgR polypeptide. An additional step of characterizing the physical, biological and/or biochemical properties of the binding partner compound is also comprehended in another embodiment of the invention. In one aspect, the NgR polypeptide/binding
5 partner complex is isolated using an antibody immunospecific for either the NgR polypeptide or the candidate binding partner compound.

In still other embodiments, either the NgR polypeptide or the candidate binding partner compound comprises a label or tag that facilitates its isolation, and methods of the invention to identify binding partner compounds include a step of isolating the NgR
10 polypeptide/binding partner complex through interaction with the label or tag. An exemplary tag of this type is a poly-histidine sequence, generally around six histidine residues, that permits isolation of a compound so labeled using nickel chelation. Other labels and tags, such as the FLAG[®] tag (Eastman Kodak, Rochester, NY), well known and routinely used in the art, are embraced by the invention.

15

(b) Immobilized NgR

In one variation of an *in vitro* assay, the invention provides a method comprising the steps of (a) contacting an immobilized NgR polypeptide, or a biologically active fragment thereof with a candidate binding partner compound and (b)
20 detecting binding of the candidate compound to the NgR polypeptide. In an alternative embodiment, the candidate binding partner compound is immobilized and binding of NgR is detected. Immobilization is accomplished using any of the methods well known in the art, including covalent bonding to a support, a bead or a chromatographic resin, as well as non-covalent, high affinity interactions such as
25 antibody binding, or use of streptavidin/biotin binding wherein the immobilized compound includes a biotin moiety. Binding of a test compound to NgR, or interaction of NgR with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge
30 tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For example, and not by way of limitation, GST-NgR fusion proteins or GST-target fusion proteins can be

- 74 -

adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or NgR protein, and the mixture is incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, and the complexes determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of NgR binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either NgR or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated NgR or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with NgR or target molecules, but which do not interfere with binding of the NgR protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or NgR trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the NgR or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the NgR or target molecule.

Detection of binding can be accomplished (i) using a radioactive label on the compound that is not immobilized, (ii) using of a fluorescent label on the non-immobilized compound, (iii) using an antibody immunospecific for the non-immobilized compound, (iv) using a label on the non-immobilized compound that excites a fluorescent support to which the immobilized compound is attached, (v) determining the activity of the NgR, as well as other techniques well known and routinely practiced in the art.

- 75 -

Determining the activity of the target molecule, for example, may be accomplished by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+} , diacylglycerol, IP_3 , etc.), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (comprising a NgR-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

(c) Competition experiments

In yet another embodiment, the assay comprises contacting the NgR protein or biologically active portion thereof with a known compound which binds NgR to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NgR protein, wherein determining the ability of the test compound to interact with a NgR protein comprises determining the ability of the test compound to preferentially bind to NgR or biologically active portion thereof as compared to the known compound.

In yet another embodiment, the cell-free assay comprises contacting the NgR protein or biologically active portion thereof with a known compound which binds NgR to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NgR protein, wherein determining the ability of the test compound to interact with a NgR protein comprises determining the ability of the NgR protein to modulate the activity of a NgR target molecule.

The cell-free assays of the present invention are amenable to use of both the soluble form or the membrane-bound form of NgR. In the case of cell-free assays comprising the membrane-bound form of NgR, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of NgR is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton[®] X-100, Triton[®] X-114, Thesit[®], Isotridecypoly(ethylene glycol ether)_n, 3-(3-cholamidopropyl)dimethylamminiol-1-propane sulfonate (CHAPS), 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-

- 76 -

propane sulfonate (CHAPSO), or N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate.

Modulators

5 Agents that modulate (*i.e.*, increase, decrease, or block) NgR activity or expression may be identified by incubating a putative modulator with a cell containing a NgR polypeptide or polynucleotide and determining the effect of the putative modulator on NgR activity or expression. The selectivity of a compound that modulates the activity of NgR can be evaluated by comparing its effects on NgR to its
10 effect on other NgR compounds. Selective modulators may include, for example, antibodies and other proteins, peptides or organic molecules which specifically bind to a NgR polypeptide or a NgR-encoding nucleic acid. Modulators of NgR activity will be therapeutically useful in treatment of diseases and physiological conditions in which normal or aberrant NgR activity is involved. NgR polynucleotides, polypeptides and
15 modulators may be used in the treatment of such diseases and conditions associated with demyelination. NgR polynucleotides and polypeptides, as well as NgR modulators, may also be used in diagnostic assays for such diseases or conditions.

Methods of the invention to identify modulators include variations on any of the methods described above to identify binding partner compounds, the variations
20 including techniques wherein a binding partner compound has been identified and the binding assay is carried out in the presence and absence of a candidate modulator. A modulator is identified in those instances where binding between the NgR polypeptide and the binding partner compound changes in the presence of the candidate modulator compared to binding in the absence of the candidate modulator compound. A
25 modulator that increases binding between the NgR polypeptide and the binding partner compound is described as an enhancer or activator, and a modulator that decreases binding between the NgR polypeptide and the binding partner compound is described as an inhibitor.

In another embodiment, modulators of NgR expression may be identified in a
30 method wherein a cell is contacted with a candidate compound and the expression of NgR mRNA or protein in the cell is determined. The level of expression of NgR mRNA or protein in the presence of the candidate compound is compared to the level

- 77 -

of expression of NgR mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of NgR expression based on this comparison. For example, when expression of NgR mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of NgR mRNA or protein expression. Alternatively, when expression of NgR mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of NgR mRNA or protein expression. The level of NgR mRNA or protein expression in the cells can be determined by methods described herein for detecting NgR mRNA or protein.

High Throughput Screening

The invention also comprehends high-throughput screening (HTS) assays to identify compounds that interact with or inhibit biological activity (*i.e.*, affect enzymatic activity, binding activity, etc.) of a NgR polypeptide. HTS assays permit screening of large numbers of compounds in an efficient manner. Cell-based HTS systems are contemplated to investigate NgR receptor-ligand interaction. HTS assays are designed to identify "hits" or "lead compounds" having the desired property, from which modifications can be designed to improve the desired property. Chemical modification of the "hit" or "lead compound" is often based on an identifiable structure/activity relationship between the "hit" and the NgR polypeptide.

Another aspect of the present invention is directed to methods of identifying compounds that bind to either NgR or nucleic acid molecules encoding NgR, comprising contacting NgR, or a nucleic acid molecule encoding the same, with a compound, and determining whether the compound binds NgR or a nucleic acid molecule encoding the same. Binding can be determined by binding assays which are well known to the skilled artisan, including, but not limited to, gel-shift assays, Western blots, radiolabeled competition assay, phage-based expression cloning, co-fractionation by chromatography, co-precipitation, cross linking, interaction trap/two-hybrid analysis, southwestern analysis, ELISA, and the like, which are described in, for example, Ausubel *et al.* (Eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, 1999, John Wiley & Sons, NY, which is incorporated herein by reference in

- 78 -

its entirety. The NgR proteins, for example, can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.*, (1993) *Cell* 72, 223-232; Madura *et al.*, (1993) *J. Biol. Chem.* 268, 12046-12054; Bartel *et al.*, (1993) *BioTechniques* 14, 920-924; Iwabuchi *et al.*, (1993) *Oncogene* 8, 1693-1696; and Brent WO 94/10300), to identify other proteins that bind to or interact with NgR ("NgR-binding proteins" or "NgR-bp") and modulate NgR activity. Such NgR-binding proteins are also likely to be involved in the propagation of signals by the NgR proteins as, for example, upstream or downstream elements of the NgR pathway.

Other assays may be used to identify specific ligands of a NgR receptor, including assays that identify ligands of the target protein through measuring direct binding of test ligands to the target protein, as well as assays that identify ligands of target proteins through affinity ultrafiltration with ion spray mass spectroscopy/HPLC methods or other physical and analytical methods. Alternatively, such binding interactions are evaluated indirectly using the yeast two-hybrid system described in Fields *et al.*, (1989) *Nature* 340, 245-246, and Fields *et al.*, (1994) *Trends Genet.* 10, 286-292, both of which are incorporated herein by reference. The two-hybrid system is a genetic assay based on the modular nature of most transcription factors used for detecting interactions between two proteins or polypeptides. It can be used to identify proteins that bind to a known protein of interest, or to delineate domains or residues critical for an interaction. Variations on this methodology have been developed to clone genes that encode DNA binding proteins, to identify peptides that bind to a protein, and to screen for drugs. The two-hybrid system exploits the ability of a pair of interacting proteins to bring a transcription activation domain into close proximity with a DNA binding domain that binds to an upstream activation sequence (UAS) of a reporter gene, and is generally performed in yeast. The assay requires the construction of two hybrid genes encoding (1) a DNA-binding domain that is fused to a first protein and (2) an activation domain fused to a second protein. The DNA-binding domain targets the first hybrid protein to the UAS of the reporter gene; however, because most proteins lack an activation domain, this DNA-binding hybrid protein does not activate transcription of the reporter gene. The second hybrid protein, which contains the activation domain, cannot by itself activate expression of the reporter gene because it does not bind the UAS. However, when both hybrid proteins are present, the

- 79 -

noncovalent interaction of the first and second proteins tethers the activation domain to the UAS, activating transcription of the reporter gene. For example, when the first protein is a NgR gene product, or fragment thereof, that is known to interact with another protein or nucleic acid, this assay can be used to detect agents that interfere with the binding interaction. Expression of the reporter gene is monitored as different test agents are added to the system. The presence of an inhibitory agent results in lack of a reporter signal. The compounds to be screened include (which may include compounds that are suspected to bind NgR, or a nucleic acid molecule encoding the same), but are not limited to, extracellular, intracellular, biological or chemical origin.

10 The function of the NgR gene product is unclear and no ligands have yet been found which bind the gene product. The yeast two-hybrid assay is useful to identify proteins that bind to the gene product. In an assay to identify proteins that bind to a NgR receptor, or fragment thereof, a fusion polynucleotide encoding both a NgR receptor (or fragment) and a UAS binding domain (*i.e.*, a first protein) may be used.

15 In addition, a large number of hybrid genes each encoding a different second protein fused to an activation domain are produced and screened in the assay. Typically, the second protein is encoded by one or more members of a total cDNA or genomic DNA fusion library, with each second protein-coding region being fused to the activation domain. This system is applicable to a wide variety of proteins, and it is not even

20 necessary to know the identity or function of the second binding protein. The system is highly sensitive and can detect interactions not revealed by other methods; even transient interactions may trigger transcription to produce a stable mRNA that can be repeatedly translated to yield the reporter protein.

Other assays may be used to search for agents that bind to the target protein.

25 One such screening method to identify direct binding of test ligands to a target protein is described in U.S. Patent No. 5,585,277, incorporated herein by reference. This method relies on the principle that proteins generally exist as a mixture of folded and unfolded states, and continually alternate between the two states. When a test ligand binds to the folded form of a target protein (*i.e.*, when the test ligand is a ligand of the target protein), the target protein molecule bound by the ligand remains in its folded

30 state. Thus, the folded target protein is present to a greater extent in the presence of a test ligand which binds the target protein, than in the absence of a ligand. Binding of

- 80 -

the ligand to the target protein can be determined by any method which distinguishes between the folded and unfolded states of the target protein. The function of the target protein need not be known in order for this assay to be performed. Virtually any agent can be assessed by this method as a test ligand, including, but not limited to, metals, polypeptides, proteins, lipids, polysaccharides, polynucleotides and small organic molecules.

Another method for identifying ligands of a target protein is described in Wieboldt *et al.* (1997) *Anal. Chem.* 69:1683-1691, incorporated herein by reference. This technique screens combinatorial libraries of 20-30 agents at a time in solution phase for binding to the target protein. Agents that bind to the target protein are separated from other library components by simple membrane washing. The specifically selected molecules that are retained on the filter are subsequently liberated from the target protein and analyzed by HPLC and pneumatically assisted electrospray (ion spray) ionization mass spectroscopy. This procedure selects library components with the greatest affinity for the target protein, and is particularly useful for small molecule libraries.

The methods of the invention also embrace ligands, especially neuropeptides, that are attached to a label, such as a radiolabel (e.g., ^{125}I , ^{35}S , ^{32}P , ^{33}P , ^3H), a fluorescence label, a chemiluminescent label, an enzymic label and an immunogenic label. Modulators falling within the scope of the invention include, but are not limited to, non-peptide molecules such as non-peptide mimetics, non-peptide allosteric effectors, and peptides. The NgR polypeptide or polynucleotide employed in such a test may either be free in solution, attached to a solid support, borne on a cell surface or located intracellularly or associated with a portion of a cell. One skilled in the art can, for example, measure the formation of complexes between NgR and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between NgR and its substrate caused by the compound being tested.

Another aspect of the present invention is directed to methods of identifying compounds which modulate (*i.e.*, increase or decrease) activity of NgR comprising contacting NgR with a compound, and determining whether the compound modifies activity of NgR. The activity in the presence of the test compared is measured to the

- 81 -

activity in the absence of the test compound. Where the activity of the sample containing the test compound is higher than the activity in the sample lacking the test compound, the compound will have increased activity. Similarly, where the activity of the sample containing the test compound is lower than the activity in the sample
5 lacking the test compound, the compound will have inhibited activity.

The present invention is particularly useful for screening compounds by using NgR in any of a variety of drug screening techniques. The compounds to be screened include (which may include compounds which are suspected to modulate NgR activity), but are not limited to, extracellular, intracellular, biologic or chemical origin.
10 The NgR polypeptide employed in such a test may be in any form, preferably, free in solution, attached to a solid support, borne on a cell surface or located intracellularly. One skilled in the art can, for example, measure the formation of complexes between NgR and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between Nogo-R and its substrate caused by the
15 compound being tested.

The activity of NgR polypeptides of the invention can be determined by, for example, examining the ability to bind or be activated by chemically synthesized peptide ligands. Alternatively, the activity of the NgR can be assayed by examining their ability to bind calcium ions, hormones, chemokines, neuropeptides,
20 neurotransmitters, nucleotides, lipids, odorants and photons. Alternatively, the activity of the NgR can be determined by examining the activity of effector molecules including, but not limited to, adenylate cyclase, phospholipases and ion channels. Thus, modulators of NgR activity may alter a NgR receptor function, such as a binding property of a receptor or an activity. In various embodiments of the method, the assay
25 may take the form of an ion flux assay, a yeast growth assay, a non-hydrolyzable GTP assay such as a [³⁵S]-GTP S assay, a cAMP assay, an inositol triphosphate assay, a diacylglycerol assay, an Aequorin assay, a Luciferase assay, a FLIPR assay for intracellular Ca²⁺ concentration, a mitogenesis assay, a MAP Kinase activity assay, an arachidonic acid release assay (e.g., using [³H]-arachidonic acid) and an assay for
30 extracellular acidification rates, as well as other binding or function-based assays of NgR activity that are generally known in the art. NgR activity can be determined by methodologies that are used to assay for FaRP activity, which is well known to those

- 82 -

skilled in the art. Biological activities of NgR receptors according to the invention include, but are not limited to, the binding of a natural or an unnatural ligand, as well as any one of the functional activities of NgRs known in the art. Non-limiting examples of NgR activities include transmembrane signaling of various forms, which
5 may involve phosphatidylinositol (PI) association and/or the exertion of an influence over PI; another exemplary activity of NgRs is the binding of accessory proteins or polypeptides that differ from known GPI proteins.

The modulators of the invention exhibit a variety of chemical structures, which can be generally grouped into non-peptide mimetics of natural NgR receptor ligands,
10 peptide and non-peptide allosteric effectors of NgR receptors, and peptides that may function as activators or inhibitors (competitive, uncompetitive and non-competitive) (e.g., antibody products) of NgR receptors. The invention does not restrict the sources for suitable modulators, which may be obtained from natural sources such as plant, animal or mineral extracts, or non-natural sources such as small molecule
15 libraries, including the products of combinatorial chemical approaches to library construction, and peptide libraries.

Other assays can be used to examine enzymatic activity including, but not limited to, photometric, radiometric, HPLC, electrochemical, and the like, which are described in, for example, ENZYME ASSAYS: A PRACTICAL APPROACH, Eisenthal and
20 Danson (Eds.), 1992, Oxford University Press, which is incorporated herein by reference in its entirety.

The use of cDNAs in drug discovery programs is well-known; assays capable of testing thousands of unknown compounds per day in high-throughput screens (HTSs) are thoroughly documented. The literature is replete with examples of the use
25 of radiolabelled ligands in HTS binding assays for drug discovery (see Williams (1991) *Med. Res. Rev.*, 11, 147-184; Sweetnam *et al.*, (1993) *J. Nat. Prod.* 56, 441-455 for review). Recombinant receptors are preferred for binding assay HTS because they allow for better specificity (higher relative purity), provide the ability to generate large amounts of receptor material, and can be used in a broad variety of formats (see
30 Hodgson (1992) *Bio/Technology* 10, 973-980; each of which is incorporated herein by reference in its entirety).

- 83 -

A variety of heterologous systems is available for functional expression of recombinant receptors that are well known to those skilled in the art. Such systems include bacteria (Strosberg *et al.* (1992) *Trends Pharmacol. Sci.* 13, 95-98), yeast (Pausch (1997) *Trends Biotechnol.* 15, 487-494), several kinds of insect cells (Vanden
5 Broeck (1996) *Int. Rev. Cytol.* 164, 189-268), amphibian cells (Jayawickreme *et al.* (1997) *Curr. Opin. Biotechnol.* 8, 629-634) and several mammalian cell lines (CHO, HEK293, COS, etc.; see Gerhardt *et al.* (1997) *Eur. J. Pharmacol.* 334, 1-23). These examples do not preclude the use of other possible cell expression systems, including cell lines obtained from nematodes (PCT application WO 98/37177).

10 In preferred embodiments of the invention, methods of screening for compounds which modulate NgR activity comprise contacting test compounds with NgR and assaying for the presence of a complex between the compound and NgR. In such assays, the ligand is typically labeled. After suitable incubation, free ligand is separated from that present in bound form, and the amount of free or uncomplexed
15 label is a measure of the ability of the particular compound to bind to NgR.

In another embodiment of the invention, high throughput screening for compounds having suitable binding affinity to NgR is employed. Briefly, large numbers of different small peptide test compounds are synthesized on a solid substrate. The peptide test compounds are contacted with NgR and washed. Bound NgR is then
20 detected by methods well known in the art. Purified polypeptides of the invention can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the protein and immobilize it on the solid support.

Generally, an expressed NgR can be used for HTS binding assays in
25 conjunction with its defined ligand. The identified peptide is labeled with a suitable radioisotope, including, but not limited to, ^{125}I , ^3H , ^{35}S or ^{32}P , by methods that are well known to those skilled in the art. Alternatively, the peptides may be labeled by well-known methods with a suitable fluorescent derivative (Baindur *et al.* (1994) *Drug Dev. Res.* 33, 373-398; Rogers (1997) *Drug Discov. Today* 2, 156-160). Radioactive
30 ligand specifically bound to the receptor in membrane preparations made from the cell line expressing the recombinant protein can be detected in HTS assays in one of several standard ways, including filtration of the receptor-ligand complex to separate

- 84 -

- bound ligand from unbound ligand (Williams (1991) *Med. Res. Rev.* 11, 147-184; Sweetnam *et al.* (1993) *J. Nat. Prod.* 56, 441-455). Alternative methods include a scintillation proximity assay (SPA) or a FlashPlate format in which such separation is unnecessary (Nakayama (1998) *Curr. Opin. Drug Disc. Dev.* 1, 85-91 Bossé *et al.* (1998) *J. Biomol. Screening* 3, 285-292). Binding of fluorescent ligands can be detected in various ways, including fluorescence energy transfer (FRET), direct spectrophotofluorometric analysis of bound ligand, or fluorescence polarization (Rogers (1997) *Drug Discov. Today* 2, 156-160; Hill (1998) *Curr. Opin. Drug Disc. Dev.* 1, 92-97).
- 10 Examples of such biological responses include, but are not limited to, the following: the ability to survive in the absence of a limiting nutrient in specifically engineered yeast cells (Pausch (1997) *Trends in Biotechnol.* 15, 487-494); changes in intracellular Ca^{2+} concentration as measured by fluorescent dyes (Murphy *et al.* (1998) *Cur. Opin. Drug Disc. Dev.* 1, 192-199). Fluorescence changes can also be used to
- 15 monitor ligand-induced changes in membrane potential or intracellular pH; an automated system suitable for HTS has been described for these purposes (Schroeder *et al.* (1996) *J. Biomol. Screening* 1, 75-80). Melanophores prepared from *Xenopus laevis* show a ligand-dependent change in pigment organization in response to heterologous NgR activation; this response is adaptable to HTS formats
- 20 (Jayawickreme *et al.* (1997) *Curr. Opin. Biotechnol.* 8, 629-634). Assays are also available for the measurement of common second messengers, including cAMP, phosphoinositides and arachidonic acid, but these are not generally preferred for HTS.
- 25 Preferred methods of HTS employing these receptors include permanently transfected CHO cells, in which agonists and antagonists can be identified by the ability to transduce the signal for the binding of Nogo in membranes prepared from these cells through the putative GPI anchor. In another embodiment of the invention, permanently transfected CHO cells could be used for the preparation of membranes which contain significant amounts of the recombinant receptor proteins; these membrane preparations would then be used in receptor binding assays, employing the
- 30 radiolabelled ligand specific for the particular receptor. Alternatively, a functional assay, such as fluorescent monitoring of ligand-induced changes in internal Ca^{2+} concentration or membrane potential in permanently transfected CHO cells containing

- 85 -

each of these receptors individually or in combination would be preferred for HTS. Equally preferred would be an alternative type of mammalian cell, such as HEK293 or COS cells, in similar formats. More preferred would be permanently transfected insect cell lines, such as *Drosophila* S2 cells. Even more preferred would be recombinant
5 yeast cells expressing the *Drosophila melanogaster* receptors in HTS formats well known to those skilled in the art (*e.g.*, Pausch (1997), above).

The invention contemplates a multitude of assays to screen and identify inhibitors of ligand binding to NgR receptors. In one example, the NgR receptor is immobilized and interaction with a binding partner is assessed in the presence and
10 absence of a candidate modulator such as an inhibitor compound. In another example, interaction between the NgR receptor and its binding partner is assessed in a solution assay, both in the presence and absence of a candidate inhibitor compound. In either assay, an inhibitor is identified as a compound that decreases binding between the NgR receptor and its binding partner. Another contemplated assay involves a variation of
15 the di-hybrid assay wherein an inhibitor of protein/protein interactions is identified by detection of a positive signal in a transformed or transfected host cell, as described in PCT publication number WO 95/20652, published August 3, 1995.

Candidate modulators contemplated by the invention include compounds selected from libraries of either potential activators or potential inhibitors. There are a
20 number of different libraries used for the identification of small molecule modulators, including: (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules. Chemical libraries consist of random chemical structures, some of which are analogs of known compounds or analogs of compounds that have been identified as "hits" or
25 "leads" in other drug discovery screens, some of which are derived from natural products, and some of which arise from non-directed synthetic organic chemistry. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms that are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of
30 plants or marine organisms. Natural product libraries include polyketides, non-ribosomal peptides, and variants (non-naturally occurring) thereof. For a review, see Cane *et al.*, *Science* (1998) 282, 63-68. Combinatorial libraries are composed of

- 86 -

large numbers of peptides, oligonucleotides, or organic compounds as a mixture. These libraries are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning, or proprietary synthetic methods. Of particular interest are non-peptide combinatorial libraries. Still other libraries of interest include peptide, 5 protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers (1997) *Curr. Opin. Biotechnol.* 8, 701-707. Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to modulate 10 activity.

Still other candidate inhibitors contemplated by the invention can be designed and include soluble forms of binding partners, as well as such binding partners as chimeric, or fusion, proteins. A "binding partner" as used herein broadly encompasses non-peptide modulators, as well as such peptide modulators as neuropeptides other 15 than natural ligands, antibodies, antibody fragments, and modified compounds comprising antibody domains that are immunospecific for the expression product of the identified NgR gene.

Other embodiments of the invention comprise using competitive screening assays in which neutralizing antibodies capable of binding a polypeptide of the 20 invention specifically compete with a test compound for binding to the polypeptide. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants with NgR. Radiolabeled competitive binding studies are described in Lin *et al.*, (1997) *Antimicrob. Agents Chemother.* 41, 2127-2131, the disclosure of which is incorporated herein by reference in its entirety.

25 In other embodiments of the invention, the polypeptides of the invention are employed as a research tool for identification, characterization and purification of interacting, regulatory proteins. Appropriate labels are incorporated into the polypeptides of the invention by various methods known in the art and the polypeptides are used to capture interacting molecules. For example, molecules are 30 incubated with the labeled polypeptides, washed to remove unbound polypeptides, and the polypeptide complex is quantified. Data obtained using different concentrations of

- 87 -

polypeptide are used to calculate values for the number, affinity, and association of polypeptide with the protein complex.

Labeled polypeptides are also useful as reagents for the purification of molecules with which the polypeptide interacts including, but not limited to, inhibitors.

- 5 In one embodiment of affinity purification, a polypeptide is covalently coupled to a chromatography column. Cells and their membranes are extracted, and various cellular subcomponents are passed over the column. Molecules bind to the column by virtue of their affinity to the polypeptide. The polypeptide-complex is recovered from the column, dissociated and the recovered molecule is subjected to protein sequencing.
- 10 This amino acid sequence is then used to identify the captured molecule or to design degenerate oligonucleotides for cloning the corresponding gene from an appropriate cDNA library.

- Alternatively, compounds may be identified which exhibit similar properties to the ligand for the NgR of the invention, but which are smaller and exhibit a longer half
- 15 time than the endogenous ligand in a human or animal body. When an organic compound is designed, a molecule according to the invention is used as a "lead" compound. The design of mimetics to known pharmaceutically active compounds is a well-known approach in the development of pharmaceuticals based on such "lead" compounds. Mimetic design, synthesis and testing are generally used to avoid
- 20 randomly screening a large number of molecules for a target property. Furthermore, structural data deriving from the analysis of the deduced amino acid sequences encoded by the DNAs of the present invention are useful to design new drugs, more specific and therefore with a higher pharmacological potency.

- Comparison of the protein sequence of the present invention with the
- 25 sequences present in all the available databases showed a significant homology with the transmembrane portion of G protein coupled receptors. Accordingly, computer modeling can be used to develop a putative tertiary structure of the proteins of the invention based on the available information of the transmembrane domain of other proteins. Thus, novel ligands based on the predicted structure of NgR can be
- 30 designed.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

Compositions and Pharmaceutical Compositions

In a particular embodiment, the novel molecules identified by the screening methods according to the invention are low molecular weight organic molecules, in which case a composition or pharmaceutical composition can be prepared thereof for oral or parenteral administration. The compositions, or pharmaceutical compositions, comprising the nucleic acid molecules, vectors, polypeptides, antibodies and compounds identified by the screening methods described herein, typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The nature of the carrier or other ingredients will depend on the specific route of administration and particular embodiment of the invention to be administered. Examples of techniques and protocols that are useful in this context are, *inter alia*, found in Remington's PHARMACEUTICAL SCIENCES, 16th ed., (1980) Osol, A (Ed.), which is incorporated herein by reference in its entirety. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, Ringer's solution, dextrose solution and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include oral and parenteral (*e.g.*, intravenous, intradermal, subcutaneous, inhalation, transdermal (topical), transmucosal and rectal administration). Solutions or suspensions used for parenteral, intradermal or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic

- 89 -

acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a NgR protein or anti-NgR antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of

- 90 -

preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They
5 can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed.
10 Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel or corn starch; a lubricant
15 such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable
20 propellant, *e.g.*, a gas such as carbon dioxide or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile
25 salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with
30 conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

- 91 -

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811. It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by any of a number of routes, *e.g.*, as described in U.S. Patent No. 5,703,055. Delivery can thus also include, *e.g.*, intravenous injection, local administration (see U.S. Patent No. 5,328,470) or stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91, 3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack or dispenser together with instructions for administration.

- 92 -

The dosage of these low molecular weight compounds will depend on the disease state or condition to be treated and other clinical factors such as weight and condition of the human or animal and the route of administration of the compound. For treating human or animals, between approximately 0.5 mg/kg of body weight to 500 mg/kg of body weight of the compound can be administered. Therapy is typically administered at lower dosages and is continued until the desired therapeutic outcome is observed.

Another aspect of the present invention is the use of the NgR nucleotide sequences disclosed herein for identifying homologs of the Nogo-R, in other animals, including but not limited to humans and other mammals and invertebrates. Any of the nucleotide sequences disclosed herein, or any portion thereof, can be used, for example, as probes to screen databases or nucleic acid libraries, such as, for example, genomic or cDNA libraries, to identify homologs using screening procedures well known to those skilled in the art. Accordingly, homologs having at least 50%, more preferably at least 60%, more preferably at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 95%, and most preferably at least 100% homology with NgR sequences can be identified.

The present compounds and methods, including nucleic acid molecules, polypeptides, antibodies, compounds identified by the screening methods described herein, have a variety of pharmaceutical applications and may be used, for example, to treat or prevent unregulated cellular growth, such as cancer cell and tumor growth. In a particular embodiment, the present molecules are used in gene therapy. For a review of gene therapy procedures, see *e.g.* Anderson *Science* (1992) 256, 808-813, which is incorporated herein by reference in its entirety.

The present invention also encompasses a method of agonizing (stimulating) or antagonizing a NgR natural binding partner associated activity in a mammal comprising administering to said mammal an agonist or antagonist to one of the above disclosed polypeptides in an amount sufficient to effect said agonism or antagonism. One embodiment of the present invention, then, is a method of treating diseases in a mammal with an agonist or antagonist of the protein of the present invention comprising administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize NgR-associated functions.

- 93 -

Methods of determining the dosages of compounds to be administered to a patient and modes of administering compounds to an organism are disclosed in U.S. Application Serial No. 08/702,282, filed August 23, 1996, and International patent publication number WO 96/22976, published August 1, 1996, both of which are
5 incorporated herein by reference in their entirety, including any drawings, figures or tables. Those skilled in the art will appreciate that such descriptions are applicable to the present invention and can be easily adapted to it.

The proper dosage depends on various factors such as the type of disease being treated, the particular composition being used and the size and physiological condition
10 of the patient. Therapeutically effective doses for the compounds described herein can be estimated initially from cell culture and animal models. For example, a dose can be formulated in animal models to achieve a circulating concentration range that initially takes into account the IC_{50} as determined in cell culture assays. The animal model data can be used to more accurately determine useful doses in humans.

15 Plasma half-life and biodistribution of the drug and metabolites in the plasma, tumors and major organs can also be determined to facilitate the selection of drugs most appropriate to inhibit a disorder. Such measurements can be carried out. For example, HPLC analysis can be performed on the plasma of animals treated with the drug and the location of radiolabeled compounds can be determined using detection
20 methods such as X-ray, CAT scan and MRI. Compounds that show potent inhibitory activity in the screening assays, but have poor pharmacokinetic characteristics, can be optimized by altering the chemical structure and retesting. In this regard, compounds displaying good pharmacokinetic characteristics can be used as a model.

Toxicity studies can also be carried out by measuring the blood cell
25 composition. For example, toxicity studies can be carried out in a suitable animal model as follows: (1) the compound is administered to mice (an untreated control mouse should also be used); (2) blood samples are periodically obtained via the tail vein from one mouse in each treatment group; and (3) the samples are analyzed for red and white blood cell counts, blood cell composition and the percent of lymphocytes
30 versus polymorphonuclear cells. A comparison of results for each dosing regime with the controls indicates if toxicity is present.

- 94 -

At the termination of each toxicity study, further studies can be carried out by sacrificing the animals (preferably, in accordance with the American Veterinary Medical Association guidelines Report of the American Veterinary Medical Assoc. Panel on Euthanasia, (1993) *J. Am. Vet. Med. Assoc.* 202:229-249). Representative
5 animals from each treatment group can then be examined by gross necropsy for immediate evidence of metastasis, unusual illness or toxicity. Gross abnormalities in tissue are noted and tissues are examined histologically. Compounds causing a reduction in body weight or blood components are less preferred, as are compounds having an adverse effect on major organs. In general, the greater the adverse effect the
10 less preferred the compound.

For the treatment of cancers the expected daily dose of a hydrophobic pharmaceutical agent is between 1 to 500 mg/day, preferably 1 to 250 mg/day, and most preferably 1 to 50 mg/day. Drugs can be delivered less frequently provided plasma levels of the active moiety are sufficient to maintain therapeutic effectiveness.
15 Plasma levels should reflect the potency of the drug. Generally, the more potent the compound the lower the plasma levels necessary to achieve efficacy.

NgR mRNA transcripts have been found in the brain and heart. SEQ ID NOs: 1 and/or, 3 will, as detailed above, enable screening the endogenous neurotransmitters/hormones/ligands which activate, agonize, or antagonize NgR and
20 for compounds with potential utility in treating disorders including CNS disorders (e.g., stroke) and degenerative disorders such as those associated with demyelination.

For example, NgR receptor activation may mediate the prevention of neurite outgrowth. Inhibition would be beneficial in both chronic and acute brain injury. See, e.g., Donovan *et al.*, (1997) *J. Neurosci.* 17, 5316-5326; Turgeon *et al.*, (1998) *J. Neurosci.* 18, 6882-6891; Smith-Swintosky *et al.*, (1997) *J. Neurochem.* 69,
25 1890-1896; Gill *et al.*, (1998) *Brain Res.* 797, 321-327; Suidan *et al.*, (1996) *Semin. Thromb. Hemost.* 22, 125-133.

Pharmacogenomics

- Agents, or modulators that have a stimulatory or inhibitory effect on NgR activity (*e.g.*, NgR gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (*e.g.*, a disease condition such as a demyelination disorder) associated with aberrant NgR activity. In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of NgR protein, expression of NgR nucleic acid or mutation content of NgR genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

- Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum (1996) *Clin. Exp. Pharmacol. Physiol.* 23, 983-985 and Linder (1997) *Clin. Chem.* 43, 254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

- As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2

- 96 -

(NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At the other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of NgR protein, expression of NgR nucleic acid, or mutation content of NgR genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a NgR modulator, such as a modulator identified by one of the exemplary screening assays described herein.

25

Monitoring Clinical Efficacy

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of NgR (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase NgR gene expression, protein levels or upregulate NgR activity, can be monitored in clinical trials of subjects exhibiting decreased NgR gene

30

- 97 -

expression, protein levels, or downregulated NgR activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease NgR gene expression, protein levels, or downregulate NgR activity, can be monitored in clinical trials of subjects exhibiting increased NgR gene expression, protein levels, or
5 upregulated NgR activity. In such clinical trials, the expression or activity of NgR and, preferably, other genes that have been implicated in, for example, a disease or disorder, can be used as a "read out" or markers of the immune responsiveness of a particular cell.

For example, genes, including NgR, that are modulated in cells by treatment
10 with an agent (*e.g.*, compound, drug or small molecule) that modulates NgR activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on demyelination disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of NgR and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a
15 gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced by one of the methods as described herein or by measuring the levels of activity of NgR or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state
20 may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, peptidomimetic, nucleic acid, small molecule, or other drug candidate
25 identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a NgR protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the NgR
30 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the NgR protein, mRNA or genomic DNA in the pre-administration sample with the NgR protein, mRNA or genomic DNA in the post

- 98 -

administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of NgR to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration
5 of the agent may be desirable to decrease expression or activity of NgR to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods
10 of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant NgR expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that
15 antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, (i) a NgR polypeptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to a NgR peptide; (iii) nucleic acids encoding a NgR peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous
20 insertion within the coding sequences of coding sequences to a NgR peptide) are utilized to "knockout" endogenous function of a NgR peptide by homologous recombination (see, *e.g.*, Capecchi (1989) *Science* 244, 1288-1292); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction
25 between a NgR peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner.
30 Therapeutics that may be utilized include, but are not limited to, a NgR peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

- 99 -

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a NgR peptide). Methods that are well-known within the art
5 include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a
10 disease or condition associated with an aberrant NgR expression or activity, by administering to the subject an agent that modulates NgR expression or at least one NgR activity. Subjects at risk for a disease that is caused or contributed to by aberrant NgR expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic
15 agent can occur prior to the manifestation of symptoms characteristic of the NgR aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of NgR aberrancy, for example, a NgR agonist or NgR antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

20 Another aspect of the invention pertains to methods of modulating NgR expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of NgR protein activity associated with the cell. An agent that modulates NgR protein activity can be an agent as described herein, such as a nucleic acid or a
25 protein, a naturally-occurring cognate ligand of a NgR protein, a peptide, a NgR peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more NgR protein activity. Examples of such stimulatory agents include active NgR protein and a nucleic acid molecule encoding NgR that has been introduced into the cell. In another embodiment, the agent inhibits one or more NgR protein activity.
30 Examples of such inhibitory agents include antisense NgR nucleic acid molecules and anti-NgR antibodies. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the

- 100 -

agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a NgR protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described
5 herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) NgR expression or activity. In another embodiment, the method involves administering a NgR protein or nucleic acid molecule as therapy to compensate for reduced or aberrant NgR expression or activity.

10 Gene Therapy

Mutations in the NgR gene that result in loss of normal function of the NgR gene product underlie NgR human disease states. The invention comprehends gene therapy to restore NgR activity to treat those disease states. Delivery of a functional NgR gene to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors,
15 and more particularly viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (*e.g.*, liposomes or chemical treatments). See, for example, Anderson (1998) *Nature*, supplement to 392(6679):25-20. For additional reviews of gene therapy technology see Friedmann (1989) *Science* 244, 1275-1281; Verma (1990) *Sci. Am.* 68-84; and Miller (1992)
20 *Nature* 357, 455-460. Alternatively, it is contemplated that in other human disease states, preventing the expression of, or inhibiting the activity of, NgR will be useful in treating disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of NgR.

The present invention provides for both prophylactic and therapeutic methods
25 of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant NgR expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that
30 antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, (i) a NgR polypeptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to a NgR

- 101 -

peptide; (iii) nucleic acids encoding a NgR peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to a NgR peptide) are utilized to "knockout" endogenous function of a NgR peptide by homologous
5 recombination (see, *e.g.*, Capecchi (1989), above); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between a NgR peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject
10 not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, a NgR peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases
15 bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a NgR peptide). Methods that are well-known within the art
20 include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a
25 disease or condition associated with an aberrant NgR expression or activity, by administering to the subject an agent that modulates NgR expression or at least one NgR activity. Subjects at risk for a disease that is caused or contributed to by aberrant NgR expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic
30 agent can occur prior to the manifestation of symptoms characteristic of the NgR aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of NgR aberrancy, for example, a NgR agonist or

- 102 -

NgR antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

Another aspect of the invention pertains to methods of modulating NgR expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of NgR protein activity associated with the cell. An agent that modulates NgR protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a NgR protein, a peptide, a NgR peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more NgR protein activity. Examples of such stimulatory agents include active NgR protein and a nucleic acid molecule encoding NgR that has been introduced into the cell. In another embodiment, the agent inhibits one or more NgR protein activity. Examples of such inhibitory agents include antisense NgR nucleic acid molecules and anti-NgR antibodies. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a NgR protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) NgR expression or activity. In another embodiment, the method involves administering a NgR protein or nucleic acid molecule as therapy to compensate for reduced or aberrant NgR expression or activity.

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figure. Such modifications are intended to fall within the scope of the appended claims.

The following Table 5 contains the sequences of exemplary polynucleotides and polypeptides of the invention.

TABLE 5

The following DNA sequence NgR2 <SEQ ID NO. 1> was identified in humans:

```

5  ATGCTGCCCCGGGCTCAGGCGCCTGCTGCAAGCTCCCGCCTCGGCCTGCCTCCTGCTGATG
   CTCCTGGCCCTGCCCCTGCGGGCCCCAGCTGCCCCATGCTCTGCACCTGCTACTCATCC
   CCGCCACCGTGAGCTGCCAGGCCAACAACTTCTCTCTGTGCCGCTGTCCCTGCCACCC
   AGCACTCAGCGACTCTTCCTGCAGAACAACTCATCCGCACGCTGCGGCCAGGCACCTTT
   GGGTCCAACCTGCTCACCCCTGTGGCTCTTCTCCAACAACCTCTCCACCATCTACCCGGGC
10  ACTTTCGCCCACTTGCAAGCCCTGGAGGAGCTGGACCTCGGTGACAACCGGCACCTGCGC
   TCGCTGGAGCCCGACACCTTCCAGGGCCTGGAGCGGCTGCAGTCGCTGCATTGTACCGC
   TGCCAGCTCAGCAGCCTGCCCGGCAACATCTTCCGAGGCCTGGTCAGCCTGCAGTACCTC
   TACCTCCAGGAGAACAGCCTGCTCCACCTACAGGATGACTTGTTCGCGGACCTGGCCAAC
   CTGAGCCACCTCTTCTCCACGGGAACCGCCTGCGGCTGCTCACAGAGCACGTGTTTCGC
15  GGCCTGGGCAGCCTGGACCGGCTGCTGCTGCACGGGAACCGGCTGCAGGGCGTGACCGC
   GCGGCCTTCCGCGGCCTCAGCCGCCTCACCATCCTCTACCTGTTCAACAACAGCCTGGCC
   TCGCTGCCCGGCGAGGCGCTCGCCGACCTGCCCTCGCTCGAGTTCCTGCGGCTCAACGCT
   AACCCTGGGCGTGCGACTGCCGCGCGCGCCGCTCTGGGCCTGGTTCCAGCGCGCGCGC
   GTGTCCAGCTCCGACGTGACCTGCGCCACCCCCCGAGCGCCAGGGCCGAGACCTGCGC
20  GCGCTCCGCGAGGCCGACTTCCAGGCGTGTCCGCCCCGCGCACCCACGCGCCGGGCGAGC
   CGCGCCCGCGGCAACAGCTCCTCAACCACTGTACGGGGTGCCGAGGCGGGGCGCCC
   CCAGCCGATCCCTCCACCCTTACCGAGATCTGCCTGCCGAAGACTCGCGGGGGCGCCAG
   GCGGGGACGCGCCTACTGAGGACGACTACTGGGGGGGCTACGGGGGTGAGGACCAGCGA
25  GGGGAGCAGATGTGCCCCGGCGCTGCCTGCCAGGCGCCCCCGACTCCCGAGGCCCTGCG
   CTCTCGGCCGGGCTCCCCAGCCCTCTGCTTGCCTCCTGCTCCTGGTGCCCCACCACCTC

```

The following amino acid sequence <SEQ ID NO. 2> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 1:

```

30  M L P G L R R L L Q A P A S A C L L L M L L A L P L A A P S C
   P M L C T C Y S S P P T V S C Q A N N F S S V P L S L P P S T
   Q R L F L Q N N L I R T L R P G T F G S N L L T L W L F S N N
   L S T I Y P G T F R H L Q A L E E L D L G D N R H L R S L E P
   D T F Q G L E R L Q S L H L Y R C Q L S S L P G N I F R G L V
   S L Q Y L Y L Q E N S L L H L Q D D L F A D L A N L S H L F L
35  H G N R L R L L T E H V F R G L G S L D R L L L H G N R L Q G
   V H R A A F R G L S R L T I L Y L F N N S L A S L P G E A L A
   D L P S L E F L R L N A N P W A C D C R A R P L W A W F Q R A
   R V S S S D V T C A T P P E R Q G R D L R A L R E A D F Q A C
   P P A A P T R P G S R A R G N S S S N H L Y G V A E A G A P P
40  A D P S T L Y R D L P A E D S R G R Q G G D A P T E D D Y W G
   G Y G G E D Q R G E Q M C P G A A C Q A P P D S R G P A L S A
   G L P S P L L C L L L L V P H H L

```

- 104 -

The following DNA sequence NgR3 <SEQ ID NO. 3> was identified in mouse:

5 ATGTCTTGGCAGTCTGGAACCACAGTGACACAATCTCCCGTGCAGGCTGCTCAGGTCTCA
GGGTGCTGTGTGGAATTGCTGCTGTTGCTGCTCGCTGGAGAGCTACCTCTGGGTGGTGGT
TGTCTCGAGACTGTGTGTGTGCTACCGTGCGCCCATGACTGTCAGCTGCCAGGCACACAAC
10 TTTGCTGCCATCCCGGAGGGCATCCCAGAGGACAGTGAGCGCATCTTCCTGCAGAACAAAT
CGCATCACCTTCTCCAGCAGGGCCACTTCAGCCCCGCCATGGTCACCCTCTGGATCTAC
TCCAACAACATCACTTTCATTGCTCCCAACACCTTCGAGGGCTTTGTGCATCTGGAGGAG
CTAGACCTTGGAGACAACCGACAGCTGCGAACGCTGGCACCCGAGACCTTCCAAGGCCTG
GTGAAGCTTCACGCCCTCTACCTCTATAAGTGTGGACTGAGCGCCCTGCCCGCAGGCATC
15 TTTGGTGGCCTGCACAGCCTGCAGTATCTCTACTTGCAGGACAACCATATCGAGTACCTC
CAAGATGACATCTTTGTGGACCTGGTCAATCTCAGTCACTTGTTTCTCCATGGTAACAAG
CTATGGAGCCTGGGCCAAGGCATCTTCGGGGCCTGGTGAACCTGGACCGGTTGCTGCTG
CATGAGAACCAGCTACAGTGGGTTCACCACAAGGCTTTCCATGACCTCCACAGGCTAACC
ACCTCTTTCTCTTCAACAACAGCCTCACTGAGCTGCAGGGTGAAGTGTCTGGCCCCCTG
20 GTGGCCTTGGAGTTCCCTTCGCCTCAATGGGAATGCTTGGGACTGTGGCTGCCGGGCACGT
TCCCTGTGGGAATGGCTGCGAAGGTTCCGTGGCTCTAGCTCTGCTGTCCCTGCGCGACC
CCCAGCTGCGGCAAGGCCAGGATCTGAAGCTGCTGAGGGTGGAGGACTTCCGGAAGTGC
ACAGGACCAAGTGTCTCCTCACCAGATCAAGTCTCACACGCTTACCACCTCTGACAGGGCT
GCCCCGAAGGAGCACCATCCGTCCCATGGGGCCTCCAGGGACAAAGGCCACCCACATGGC
25 CATCCGCCTGGCTCCAGGTCAGGTTACAAGAAGGCAGGCAAGAACTGCACCAAGCCACAGG
AACCAGAACAGATCTCTAAGGTGAGCTCTGGGAAAGAGCTTACCGAACTGCAGGACTAT
GCCCCGACTATCAGCACAAGTTCAGCTTTGACATCATGCCACCGCACGACCCAAGAGG
AAGGGCAAGTGTGCTCGCAGGACCCCATCCGTGCCCCCAGTGGGGTGCAGCAGGCATCC
TCAGGCACGGCCCTTGGGGCCCCACTCTGGCCTGGATACTGGGGCTGGCAGTCACTCTC
CGC

The following protein sequence <SEQ ID NO. 4> is deduced protein of SEQ ID
NO:3:

30 MSWQSGTTVTQSPVQAAQVSGCCVELLLLLL
AGELPLGGGCPRDCVCYPAPMTVSCQAHNFA
AIPEGIPEDSERIFLQNNRITFLQQGHFSPA
MVTLWIYSNNITFIAPNTFEGFVHLEELDLG
DNRQLRTLAPETFQGLVKLHALYLYKCLSA
LPAGIFGGLHSLQYLYLQDNHIEYLQDDIFV
35 DLVNLSHLFLHGNKLWSLQGQIFRGLVNLDR
LLLHENQLQWVHHKAFHDLHRLTTLFLFNNS
LTELQGDCLAPLVALEFLRLNGNAWDCGCRA
RSLWEWLRFRFRGSSSAVPCATPELRQGQDLK
LLRVEDFRNCTGPVSPHQIKSHTLTTSDDRAA
40 RKEHHPSHGASRDKGHPHGHPPGSRSGYKKA
GKNCTSHRNRNQISKVSSGKELTELQDYAPD
YQHKFSFDIMPTARPKRKGKCARRTPIRAPS
GVQQASSGTALGAPLLAWILGLAVTLR

- 105 -

The following protein sequence <SEQ ID NO. 5> is NgR1 from humans:

MKRASAGGSRL LAWVLWLQAWQVAAPCPGA
 C
 5 CYNEPKVTTSCPQQGLQAVPVGIPAASQRI
 FLHG NRISHVPAASFRA CRNL TILWLHSNVL
 ARIDAA AFTGLALLEQLDLSDNAQLRSVDPA
 TFHGLGRLHTLHLDR CGLQELGPGLFRGLAA
 10 LQYLYLQDNALQALPDDTFRDLGNLTHLFLH
 GNRIS SVPERAFRGLHSLDRLLLHQNRVAHV
 HPHAFRDLGRLMTLYLFANNLSALPTEALAP
 LRALQYLRLNDNPWVCD CRARPLWAWLQKFR
 GSSE VPCSLPQRLAGRDLKRLAANDLQGCA
 VATGPYHPIWTGRATDEEPLGLPKCCQPDAA
 15 DKASVLEPGRPASAGNALKGRVPPGDSPPGN
 GSGPRHINDSPFGTLPGSAEPPLTAVRPEG
 EPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
 GSGGGGTGDSESGALPSLTCSLTPLGLALV
 LWTVLGPC

The following amino acid sequence <SEQ ID NO:6> is a Consensus Sequence of
 20 NgR based on homology with NgR1

C P X X C X C Y X X P X X T X S C X X X X X X X X P X
 X X P X X X X R X F L X X N X I X X X X X X F X X X X X X X L W X
 X S N X X X X I X X X X F X X X X X L E X L D L X D N X X L R
 25 X X X P X T F X G L X X L X L X L X X C X L X X L X X X F X
 G L X X L Q Y L Y L Q X N X X X X L X D D X F X D L X N L X H
 L F L H G N X X X X X X X X X F R G L X X L D R L L L H X N X
 X X X V H X X A F X X L X R L X X L X L F X N X L X X L X X X
 X L A X L X X L X X L R L N X N X W X C X C R A R X L W X W X
 30 X X X R X S S X V X C X X P X X X X G X D L X X L X X X D X
 X X C X X X X P X X P X X X X X X X X X X X X X X X X X X
 X X X X X X X X X X X X X X X X G X X X X X X X X X X X X
 P P X X X S X
 X
 35 X L X X X X X
 X X X X X L

The following protein sequence <SEQ ID NO:7> is the 66 amino acid active domain
 of Nogo:

40 R I Y K G V I Q A I Q K S D E G H P F R A Y L E S E V A I S E
 E L V Q K Y S N S A L G H V N C T I K E L R R L F L V D D L V
 D S L K

The following protein sequence <SEQ ID NO:8> is the amino acid sequence of the mature NgR2:

5 CPMLCTCYSSPPTVSCQANNFSSVPLSLPPS
TQRLFLQNNLIRTLRPGTFGSNLLTLWLFSN
NLSTIYPGTFRHLQALEELDLGDNRHLSLE
PDTFQGLERLQSLHLYRCQLSSSLPGNIFRGL
VSLQYLYLQENSLHLQDDLFADLANLSHLF
LHGNRLRLLTEHVFRGLGSLDRLLLHGNRLQ
10 GVHRAAFRGLSRLTILYLFNNSLASLPGEAL
ADLPSLEFLRLNANPWACDCRARPLWAWFQR
ARVSSSDVTCATPPERQGRDLRALREADFQA
CPPAAPTRPGSRARGNSSSNHLYGVAEAGAP
PADPTLYRDLPAEDSRGRQGGDAPTEDDYW
15 GGYGGEDQRGEQMCPGAACQAPPDSRGPALS
AGLPSPLLCLLLLVP HHL

The following protein sequence <SEQ ID NO:9> is the amino acid sequence of the mature NgR3:

20 CPRDCVCYPAPMTVSCQAHNFAAIPEGIPED
SERIFLQNNRITFLQQGHFSPAMVTLWIYSN
NITFIAPNTFEGFVHLEELDLGDNRLRLTLA
PETFQGLVKLHALYLYKCGLSALPAGIFGGL
HSLQYLYLQDNHIEYLQDDIFVDLVNLSHLF
LHGNKLWSLGQGIFRGLVNLDRLLLHENQLQ
25 WVHHKAFHDLHRLTTLFLFNNSLTELGDCCL
APLVALEFLRLNGNAWDCGCRARSLWEWLR
FRGSSSAVPCATPELRQGQDLKLLRVEDFRN
CTGPVSPHQIKSHTLTTSDDRAARKEHHP SHG
ASRDKGHPHGHPPGSRSGYKKAGKNCTSHRN
30 RNQISKVSSGKELTELQDYAPDYQHKFSFDI
MPTARPKRKGGKCARRTPIRAPSGVQQASSGT
ALGAPLLAWILGLAVTLR

35 The following amino acid sequence <SEQ ID NO:10> is a conserved cysteine motif
(Cysteine domain 1) of the NgR and homologs based on the Consensus Sequence:

CPXXCXCXYPXXTXSC

- 107 -

The following amino acid sequence <SEQ ID NO:11> is a conserved cysteine motif (Cysteine domain 2) of the NgR and homologs based on the Consensus Sequence:

N X W X C X C R A R X L W X W X X X R X S S S X V X C X X P
X X X X G X D L X X L X X X D X X X C

5

The following amino acid sequence <SEQ ID NO:12> is a conserved Leucine-rich domain of the NgR and homologs based on the Consensus Sequence:

R X F L X X N X I X X X X X X F X X X X X X X L W X X S N
X X X X I X X X F X X X X X L E X L D L X D N X X L R X X X
P X T F X G L X X L X L X L X X C X L X X L X X X X F X G L X
X L Q Y L Y L Q X N X X X X L X D D X F X D L X N L X H L F L
H G N X X X X X X X X F R G L X X L D R L L H X N X X X X
V H X X A F X X L X R L X X L X L F X N X L X X L X X X X L A
X L X X L X X L R L

10

15

Unless otherwise indicated, X is any amino acid. For example, X where indicated may be no amino acid. Additional features of the invention will be apparent from the following Examples. Examples 1-5 are actual, while the remaining Examples are prophetic.

20

As shown by the following Examples, a gene encoding novel NgRs have been identified by computational analysis of DNA sequence data. The proteins encoded by NgR2 and NgR3 have a putative signal sequence, eight leucine-rich repeat domains in a conserved leucine-rich region (SEQ ID NO:12), a conserved cysteine-rich region (SEQ ID NO:10) N-terminal to the leucine-rich region, a second cysteine-rich domain (SEQ ID NO:11) C-terminal to the leucine-rich region, and a putative glycosylphosphatidylinositol-linkage (GPI-linkage) site. NgR2 and NgR3 differ from the previously identified NgR sequence. The NgR homologs, when compared to known NgRs, show a consensus sequence (SEQ ID NOs:6). The putative mature NgR2 and NgR3 are shown in Table 5 as SEQ ID NOs: 8 and 9, respectively.

25

30

Example 1: Tblastn query of the HTG database

The protein sequence for the human NgR (NgR1) (SEQ ID NO:5) was used to query the high throughput genomic (HTG) database the use of which is familiar to those skilled in the art. The HTG database is a part of GenBank, a comprehensive

35

- 108 -

NIH genetic sequence database, which includes an annotated collection of all publicly available DNA sequences (*Nucleic Acids Res.* (2000) 28, 15-8). The HTG database includes sequences obtained from genomic DNA. Within genomic DNA, genes are typically encoded by multiple segments of DNA called exons. Thus when one aligns a
5 cDNA sequence (or a protein sequence encoded by a cDNA sequence) to a genomic sequence, the sequence will be broken up into segments depending on the number of exons in the gene.

The BLAST algorithm, which stands for Basic Local Alignment Search Tool is suitable for determining sequence similarity (Altschul *et al.*, (1990) *J. Mol. Biol.* 215,
10 403-410, which is incorporated herein by reference in its entirety). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). The basic BLAST algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some
15 positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased.
20 Extension for the word hits in each direction are halted when: 1) the cumulative alignment score falls off by the quantity X from its maximum achieved value; 2) the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or 3) the end of either sequence is reached. The Blast algorithm parameters W, T and X determine the sensitivity and speed of the
25 alignment. The Blast program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff *et al.*, (1992) *Proc. Natl. Acad. Sci. USA* 89, 10915-10919, which is incorporated herein by reference in its entirety) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands.

The BLAST algorithm (Karlin *et al.*, (1993) *Proc. Natl. Acad. Sci. USA* 90,
30 5873-5787, which is incorporated herein by reference) and Gapped BLAST perform a statistical analysis of the similarity between two sequences. One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which

- 109 -

provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a NgR gene or cDNA if the smallest sum probability in comparison of the test nucleic acid to a NgR nucleic acid is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

To query the HTG database with the NgR protein sequence, we used a variation of the BLAST algorithm known as the tblastn program, which compares a protein query sequence against a nucleotide sequence database dynamically translated in all reading frames (*J. Mol. Biol.* (1990) 215, 403-410; *Nucleic Acids Res.* (1997) 25, 3389-3402). The results of the tblastn search indicated the presence of genes in the database with a significant identity to the NgR. In addition to finding hits to genomic clones which contain the human and mouse NgR genes, we found hits to clones where the identity was not as high, but still very significant. Three human clones were found (Accession numbers: AC068514, AC016869, AC013606) with an e-value of $4e-43$ and one mouse clone was found (Accession No. AC021768) with an e-value of $1e-78$. The three human clones all appeared to encode the same gene, so further analysis was confined to AC013606.

20 **Example 2: Prediction of the human NgR2 protein sequence (AC013606)**

The human NgR protein sequence aligned with two regions of translated sequence from nucleotide sequence AC013606, indicating that the new gene was encoded by at least two exons. In order to define the complete gene, we used the computer program GENSCAN[™] (*J. Mol. Biol.* (1997) 268, 78-94) which can identify complete exon/intron structures of genes in genomic DNA. The gene prediction by GENSCAN[™] contained seven exons. By comparing these predicted exons to the NgR, it was concluded that the new human gene contains two of these exons and a part of another (containing the initiating methionine). The predicted cDNA (mRNA) encoded by these three exons was assembled from AC013606 (HTG11; deposited March 2000; length = 143899; GenBank release 118.0; SEQ ID NO: 15) by combining nucleotides from the three exons whose coordinates are: 123292-123322 (exon 1); 130035-130516 (exon 2); and 138589-139335 (exon 3). The sequence for this cDNA

- 110 -

sequence is SEQ ID NO:1 (nucleotide sequence of human NgR2; AC013606). The translation of this cDNA provides the protein sequence of human NgR2 (SEQ ID NO:2).

We used the protein sequence of human NgR2 as a query sequence against the human EST database. A number of hits of high significance were found indicating that the NgR2 mRNA is expressed in a number of tissues including fetal brain. Furthermore, two of these ESTs provided support for the exon structure that we deduced. One EST (Accession No: GB_EST19:AI346757) contains 565 nucleotides corresponding to amino acids 84-271 of the human NgR2 (SEQ ID No:4). This spans the second intron located between amino acids 171 and 172, and provides positive evidence for the splicing of exons 2 and 3 at the mRNA level. Another EST (GB_EST26:AI929019) contains 545 nucleotides, part of which corresponds to amino acids 1-75 of the human NgR2 (SEQ ID NO:2). This spans the first intron located between amino acids 10 and 11, and provides positive evidence for the splicing of exons 1 and 2 at the mRNA level.

Example 3: Prediction of the mouse NgR3 protein sequence (AC021768)

The human NgR protein sequence aligned with only one region of translated sequence from nucleotide sequence AC021768, indicating that most of the new mouse gene was encoded by one large exon. However, upon inspection, the protein encoded by this exon was missing an initiating methionine. In order to define the complete gene, we used the computer program GENSCAN as described above. The gene prediction by GENSCAN contained two exons; the large one found by visual inspection and a short one at the 5' end which provided an initiating methionine. The predicted cDNA (mRNA) encoded by these two exons was assembled from AC021768 (HTG14; deposited March 2000; length = 215980; GenBank release 118.0; SEQ ID NO: 16) by combining nucleotides from the two exons whose coordinates are: the complement of 164265-164325 (exon 1); and the complement of 155671-156992 (exon 2). The sequence for this cDNA sequence is SEQ ID NO:3 (nucleotide sequence of mouse NgR3; AC021768). The translation of this cDNA provides the protein sequence of mouse NgR3 (SEQ ID NO:4).

- 111 -

We used the protein sequence of mouse NgR3 as a query sequence against the mouse EST database. One hit of high significance was found indicating that the NgR2 mRNA is expressed in the heart. This EST (GB_EST20:AI428334) contains 463 nucleotides, part of which correspond to amino acids 45-193 of mouse NgR3 (SEQ ID NO:4).

Example 4: Similarity between the NgRs

An alignment between NgR1 and the two new receptors is shown in Fig. 1A-1B. The similarities between these proteins include:

- 10 (1) The SignalP program, which locates the signal sequence cleavage position, predicts a cleavage before the first conserved cysteine in all the proteins. Thus the mature protein in all cases will have a cysteine at the N-terminus.
- 15 (2) All proteins contain eight Leucine Rich Repeats (LRR). LRRs are short sequence motifs present in a number of proteins with diverse functions and cellular locations. These repeats are usually involved in protein-protein interactions. Each LRR is composed of a beta-alpha unit.
- 20 (3) All three proteins contain a leucine rich repeat N-terminal domain (LRRNT), in which four cysteines are conserved. LRRs are often flanked by cysteine rich domains at both their N and C termini.
- (4) All three proteins contain a LRR C-terminal domain (LRRCT). The LRRCTs of the three NgR proteins can be distinguished from those of other LRR containing proteins, by the pattern of typtophans and cysteines which are completely conserved in this domain.
- 25 (5) All three proteins contain a conserved cysteine in the fourth LRR domain.
- (6) All three proteins contain a conserved potential glycosylation site in the eighth LRR domain.
- 30 (7) NgR2 and NgR3 have a hydrophobic C-terminus, as does NgR1, an indication that they probably also undergo a modification similar to NgR1, where a GPI moiety is covalently linked to a C-terminal amino acid. This allows the protein to remain tethered to the cell.

- 112 -

Example 5: Preparation of Nogo Proteins

A Nogo binding assay was developed which utilizes a method widely used in examining semaphorin and ephrin axonal guidance function (Flanagan & Vanderhaeghen (1998) *Annu. Rev. Neurosci.* 21,3 09-345; Takahashi *et al.*, (1999) *Cell* 99, 59-69). It involves fusing a secreted placental alkaline phosphatase (AP) moiety to the ligand in question to provide a biologically active receptor binding agent which can be detected with an extremely sensitive colorimetric assay. For Nogo, an expression vector is created encoding a signal peptide, a His6 tag for purification, AP, and the 66 amino acid active domain of Nogo. The fusion protein can be purified from the conditioned medium of transfected cells in milligram amounts. This protein is biologically active as a growth cone collapsing agent with an EC_{50} of 1 nM.

Alternatively, a glutathione-S-transferase Nogo (GST-Nogo) fusion protein may be prepared. For GST-Nogo, an expression vector (*e.g.*, a pGEX vector) is created encoding a signal peptide, GST, and the 66 amino acid active domain of Nogo. GST-Nogo may be purified from the culture medium and used as a GST fusion protein, or GST may be cleaved from the Nogo portion of the fusion protein with an enzyme that recognizes the specific amino acid cleavage site engineered between the GST portion and the Nogo portion of the fusion protein. Such sites are part of the commercially available GST vectors. The specific cleavage sites and enzymes may be used in accordance with the Manufacturer's specifications.

It has been found that AP-Nogo is actually slightly more potent than GST-Nogo, perhaps because the protein is synthesized in a eukaryotic rather than a prokaryotic cell.

Binding of Nogo to immobilized NgR homologs may be performed in an ELISA-type assay in which AP-Nogo is allowed to react with an immobilized receptor homolog. Specificity of binding may be demonstrated in a competitive binding assay using increasing amounts of GST-Nogo in the type of assay to show a decreasing amount of binding of AP-Nogo (as judged in the colorimetric assay).

Example 6: Transfected COS Cell binding Assays

The homologs of the present invention may be used in transfection studies in COS cells to demonstrate binding of Nogo. Specifically, nucleotide sequences

- 113 -

encoding NgR2 and NgR3 may be transfected into COS cells using a suitable vector. Non-transfected COS-7 cells do not bind AP-Nogo. However, transfection of COS cells with nucleic acid sequences encoding NgRs will make them capable of binding Nogo. AP alone does not bind with any stable affinity to these transfected cells,
5 indicating that any affinity of Nogo for NgR2 or NgR3 would be due to the 66 amino acids derived from Nogo. Furthermore, specific affinity of Nogo for the NgR2 or NgR3 proteins may be tested in displacement of AP-Nogo assays using GST-Nogo. NgR2 and/or NgR3 may also bind homologs of Nogo, which may also be tested using this assay.

10

Example 7: Expression of NgR in Human Cell Lines using Northern Blot and a Random-Primed Probe

A Northern blot is purchased from a commercial source, or RNA samples from cells of interest are run on an agarose gel and blotted to a membrane using any of the
15 well known techniques for Northern blotting. The blot is probed with a fragment of NgR2 (SEQ ID NO:1) or NgR3 (SEQ ID NO:3). The probe is prepared from 50 ng of cDNA labeled by a random-primed method (Feinberg and Vogelstein (1983) *Anal. Biochem.* 132, 6-13). Hybridization is carried out at 68°C for 1 hour in ExpressHyb™ solution (Clontech, Cat. No. 8015-1) followed by washing with 2X
20 SSC/0.05% SDS at room temperature and two washes with 0.1X SSC/0.1% SDS at 50°C. Expression of NgR2 and/or NgR3 can be assessed by the presence of an appropriately sized band on the blot.

25

Example 8: Cloning of cDNA corresponding to NgRs

To obtain the full-length clone corresponding to NgR2 from a cDNA library, the following method may be used. A cDNA library is generated using standard methods from a tissue known to contain NgR2. Such a tissue was identified in Example 2. 1×10^6 plaque forming units from the cDNA library may be screened in
30 duplicate on OPTITRAN™ filters. The filters are hybridized with ³²P-labeled oligonucleotides that are generated from the ESTs corresponding to portions of NgR2. The hybridization reaction may consist of 400 mls plaque screen buffer (50mM Tris pH

- 114 -

7.5, 1M NaCl, 0.1% Sodium pyrophosphate, 0.2% Polyvinylpyrrolidone and 0.2% Ficoll) containing 10% Dextran sulfate and 100µg/ml tRNA and 80 pmol each ³²P-labeled oligonucleotide at 65°C overnight. The filters are washed twice with 2X SSC/1%SDS and twice with 1X SSC/1%SDS and exposed to film. Duplicate positives
5 are purified. DNA from each of these clones is analyzed by restriction enzyme digest followed by agarose gel electrophoresis and Southern blotting. The filters are hybridized to the ³²P-labeled oligonucleotides used for the original hybridization to confirm that inserts hybridize to the probe. The insert is then sequenced to confirm that it represents the cDNA for NgR2. Similar methods may be used to generate a
10 full-length clone corresponding to NgR3.

Alternatively, a full-length clone of NgR2 or NgR3 can be obtained by a person of ordinary skill in the art employing conventional PCR techniques.

Example 9: Hybridization Analysis to demonstrate NgR expression in the brain

15 The expression of NgR in mammals, such as the rat, may be investigated by *in situ* hybridization histochemistry. To investigate expression in the brain, for example, coronal and sagittal rat brain cryosections (20 µm thick) are prepared using a Reichert-Jung cryostat. Individual sections are thaw-mounted onto silanized, nuclease-free slides (CEL Associates, Inc., Houston, TX), and stored at -80°C. Sections are
20 processed starting with post-fixation in cold 4% paraformaldehyde, rinsed in cold phosphate-buffered saline (PBS), acetylated using acetic anhydride in triethanolamine buffer, and dehydrated through a series of alcohol washes in 70%, 95%, and 100% alcohol at room temperature. Subsequently, sections are delipidated in chloroform, followed by rehydration through successive exposure to 100% and 95% alcohol at
25 room temperature. Microscope slides containing processed cryosections are allowed to air dry prior to hybridization. Other tissues may be assayed in a similar fashion.

A NgR-specific probe may be generated using PCR. Following PCR amplification, the fragment is digested with restriction enzymes and cloned into pBluescript II cleaved with the same enzymes. For production of a probe specific for
30 the sense strand of NgR, a cloned NgR fragment cloned in pBluescript II may be linearized with a suitable restriction enzyme, which provides a substrate for labeled run-off transcripts (*i.e.*, cRNA riboprobes) using the vector-borne T7 promoter and

- 115 -

commercially available T7 RNA polymerase. A probe specific for the antisense strand of NgR may also be readily prepared using the NgR clone in pBluescript II by cleaving the recombinant plasmid with a suitable restriction enzyme to generate a linearized substrate for the production of labeled run-off cRNA transcripts using the T3 promoter and cognate polymerase. The riboprobes may be labeled with [³⁵S]-UTP to yield a specific activity of about 0.40×10^6 cpm/pmol for antisense riboprobes and about 0.65×10^6 cpm/pmol for sense-strand riboprobes. Each riboprobe may be subsequently denatured and added (2 pmol/ml) to hybridization buffer which contains 50% formamide, 10% dextran, 0.3 M NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA, 1X Denhardt's Solution, and 10 mM dithiothreitol. Microscope slides containing sequential brain cryosections may be independently exposed to 45 μ l of hybridization solution per slide and silanized cover slips may be placed over the sections being exposed to hybridization solution. Sections are incubated overnight (15-18 hours) at 52°C to allow hybridization to occur. Equivalent series of cryosections are then exposed to sense or antisense NgR-specific cRNA riboprobes.

Following the hybridization period, coverslips are washed off the slides in 1X SSC, followed by RNase A treatment involving the exposure of slides to 20 μ g/ml RNase A in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.5 M EDTA, and 0.5 M NaCl for 45 minutes at 37°C. The cryosections are then subjected to three high-stringency washes in 0.1 X SSC at 52°C for 20 minutes each. Following the series of washes, cryosections are dehydrated by consecutive exposure to 70%, 95%, and 100% ammonium acetate in alcohol, followed by air drying and exposure to Kodak BioMax™ MR-1 film. After 13 days of exposure, the film is developed, and any significant hybridization signal is detected. Based on these results, slides containing tissue that hybridized, as shown by film autoradiograms, are coated with Kodak NTB-2 nuclear track emulsion and the slides are stored in the dark for 32 days. The slides are then developed and counterstained with hematoxylin. Emulsion-coated sections are analyzed microscopically to determine the specificity of labeling. The signal is determined to be specific if autoradiographic grains (generated by antisense probe hybridization) are clearly associated with cresyl violet-stained cell bodies. Autoradiographic grains found between cell bodies indicate non-specific binding of the probe.

- 116 -

In some cases, such as using a probe to detect a NgR homolog in a heterologous species, in order to achieve optimal hybridization, it may be necessary to decrease the stringency conditions. Such conditions are well known to those of ordinary skill in the art and examples are provided above.

5 Expression of NgR in the brain provides an indication that modulators of NgR activity have utility for treating neurological disorders. Some other diseases for which modulators of NgR may have utility include depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Use of NgR modulators, including NgR ligands and anti-NgR antibodies, to treat individuals having
10 such disease states is intended as an aspect of the invention.

Example 10: Northern Blot Analysis of NgR-RNA with a PCR-generated Probe

Northern blot hybridizations may be performed to examine the expression of NgR mRNA. A clone containing at least a portion of the sequence of SEQ ID NO:1
15 may be used as a probe. Vector-specific primers are used in PCR to generate a hybridization probe fragment for ³²P-labeling. The PCR is performed as follows:

	Mix:	1μl	NgR-containing plasmid
		2μl	fwd primer (10-50 pM)
20		2μl	rev primer (10-50 pM)
		10μl	10xPCR buffer (such as that provided with the enzyme, Amersham Pharmacia Biotech)
		1μl	10mM dNTP (such as #1 969 064 from Boehringer Mannheim)
		0.5μl	Taq polymerase (such as #27-0799-62, Amersham Pharmacia Biotech)
25		83.5μl	water

- 117 -

PCR is performed in a Thermocycler using the following program:

	94°C	5min		
5	94°C	1min		30 cycles
	55°C	1min		
	72°C	1min		
10	72°C	10min		

- The PCR product may be purified using QIAquick PCR Purification Kit (#28104) from Qiagen, and radiactively labeled with 32 P-dCTP (#AA0005/250, Amersham Pharmacia Biotech) may be done by random priming using "Ready-to-go DNA Labeling Beads" (#27-9240-01) from Amersham Pharmacia Biotech.
- Hybridization is carried out on Human Multiple Tissue Northern Blot from Clontech as described in manufacturer's protocol, or on a Northern Blot prepared by running RNA samples from cells of interest on an agarose gel and blotting to a membrane using any of the known Northern blotting protocols. After exposure overnight on Molecular Dynamics Phosphor Imager screen (#MD146-814) bands of an appropriate size are visualized.

Example 11: Recombinant Expression of NgR in Eukaryotic Host Cells

A. Expression of NgR in Mammalian Cells

- To produce NgR protein, a NgR-encoding polynucleotide is expressed in a suitable host cell using a suitable expression vector and standard genetic engineering techniques. For example, a NgR-encoding sequence described in Table 4 is subcloned into the commercial expression vector pzeoSV2 (Invitrogen, San Diego, CA) and transfected into Chinese Hamster Ovary (CHO) cells using the transfection reagent FuGENE6™ (Boehringer-Mannheim) and the transfection protocol provided in the product insert. Other eukaryotic cell lines, including human embryonic kidney (HEK 293) and COS cells, are suitable as well. Cells stably expressing NgR are selected by growth in the presence of 100 µg/ml zeocin (Stratagene, LaJolla, CA). As an alternative to FuGENE6™, the expression vector may carry the gene for dihydrofolate reductase (dhfr) and selection of clones with methotrexate (MTX) drug pressure

- 118 -

allows for stable transformation of CHO cells. Optionally, NgR may be purified from the cells using standard chromatographic techniques. To facilitate purification, antisera is raised against one or more synthetic peptide sequences that correspond to portions of the NgR amino acid sequence, and the antisera is used to affinity purify Nogo-R.

- 5 The NgR also may be expressed in-frame with a tag sequence (*e.g.*, polyhistidine, hemagglutinin, FLAG) to facilitate purification. Moreover, it will be appreciated that many of the uses for NgR polypeptides, such as assays described below, do not require purification of NgR from the host cell.

10 **B. Expression of NgR in CHO cells**

- For expression of NgR in Chinese hamster ovary (CHO) cells, a plasmid bearing the relevant NgR coding sequence is prepared, using a vector which also bears the selectable marker dihydrofolate reductase (DHFR). The plasmid is transfected into CHO cells. Selection under MTX drug pressure allows for preparation of stable
- 15 transformants of a NgR (NgR2 or NgR3) in an expression plasmid carrying a selectable marker such as DHFR.

C. Expression of NgR in 293 cells

- For expression of NgR in mammalian cells 293 (transformed human, primary
- 20 embryonic kidney cells), a plasmid bearing the relevant NgR coding sequence is prepared, using vector pSecTag2A (Invitrogen). Vector pSecTag2A contains the murine IgK chain leader sequence for secretion, the c-myc epitope for detection of the recombinant protein with the anti-myc antibody, a C-terminal polyhistidine for purification with nickel chelate chromatography, and a Zeocin resistant gene for
- 25 selection of stable transfectants. The forward primer for amplification of this NgR cDNA is determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce the HindIII cloning site and nucleotides matching the NgR sequence. The reverse primer is also determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce an *XhoI* restriction site for cloning
- 30 and nucleotides corresponding to the reverse complement of the NgR sequence. The PCR conditions are 55°C as the annealing temperature. The PCR product is gel purified and cloned into the *HindIII-XhoI* sites of the vector.

- 119 -

The DNA is purified using Qiagen chromatography columns and transfected into 293 cells using DOTAP™ transfection media (Boehringer Mannheim, Indianapolis, IN). Transiently transfected cells are tested for expression after 24 hours of transfection, using western blots probed with anti-His and anti-NgR peptide antibodies. Permanently transfected cells are selected with Zeocin and propagated. Production of the recombinant protein is detected from both cells and media by Western blots probed with anti-His, anti-Myc or anti-NgR peptide antibodies.

D. Transient Expression of Nogo-R in COS cells

For expression of the NgR in COS7 cells, a polynucleotide molecule having a nucleotide sequence of SEQ ID NO:1, for example, can be cloned into vector p3-CI. This vector is a pUC18-derived plasmid that contains the HCMV (human cytomegalovirus) promoter-intron located upstream from the bGH (bovine growth hormone) polyadenylation sequence and a multiple cloning site.

The forward primer is determined by routine procedures and preferably contains a 5' extension which introduces an *Xba*I restriction site for cloning, followed by nucleotides which correspond to a nucleotide sequence of SEQ ID NO:1. The reverse primer is also determined by routine procedures and preferably contains 5'-extension of nucleotides which introduces a *Sa*I cloning site followed by nucleotides which correspond to the reverse complement of a nucleotide sequence of SEQ ID NO:1.

The PCR consists of an initial denaturation step of 5 min at 95°C, 30 cycles of 30 sec denaturation at 95°C, 30 sec annealing at 58°C and 30 sec extension at 72°C, followed by 5 min extension at 72°C. The PCR product is gel purified and ligated into the *Xba*I and *Sa*I sites of vector p3-CI. This construct is transformed into *E. coli* cells for amplification and DNA purification. The DNA is purified with Qiagen chromatography columns and transfected into COS 7 cells using Lipofectamine™ reagent from BRL, following the manufacturer's protocols. Forty-eight and 72 hours after transfection, the media and the cells are tested for recombinant protein expression.

NgR expressed from a COS cell culture can be purified by concentrating the cell-growth media to about 10 mg of protein/ml, and purifying the protein by, for

- 120 -

example, chromatography. Purified NgR is concentrated to 0.5 mg/ml in an Amicon concentrator fitted with a YM-10 membrane and stored at -80°C. NgR3 may also be expressed using this method and the nucleotide sequence of SEQ ID NO:3 or SEQ ID NO:13.

5

E. Expression of NgR in Insect Cells

For expression of NgR in a baculovirus system, a polynucleotide molecule having a nucleotide sequence of SEQ ID NO:1, 3 or 13 can be amplified by PCR. The forward primer is determined by routine procedures and preferably contains a 5' extension which adds the *NdeI* cloning site, followed by nucleotides which correspond to a nucleotide sequence of SEQ ID NO:1 (or SEQ ID NO:3 or SEQ ID NO:13, respectively). The reverse primer is also determined by routine procedures and preferably contains a 5' extension which introduces the *KpnI* cloning site, followed by nucleotides which correspond to the reverse complement of a nucleotide sequence of SEQ ID NO:1 (or SEQ ID NO:3 or SEQ ID NO:13, respectively).

The PCR product is gel purified, digested with *NdeI* and *KpnI*, and cloned into the corresponding sites of vector pACHTL-A (Pharmingen, San Diego, CA). The pAcHTL expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV), and a 6XHis tag upstream from the multiple cloning site. A protein kinase site for phosphorylation and a thrombin site for excision of the recombinant protein precede the multiple cloning site is also present. Of course, many other baculovirus vectors could be used in place of pAcHTL-A, such as pAc373, pVL941 and pAcIM1. Other suitable vectors for the expression of NgR polypeptides can be used, provided that the vector construct includes appropriately located signals for transcription, translation, and trafficking, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in Luckow *et al.*, Virology 170:31-39, among others.

The virus is grown and isolated using standard baculovirus expression methods, such as those described in Summers *et al.* (1987) A MANUAL OF METHODS FOR BACULOVIRUS VECTORS AND INSECT CELL CULTURE PROCEDURES, Texas Agricultural Experimental Station Bulletin No. 1555.

- 121 -

In a preferred embodiment, pAcHLT-A containing NgR gene is introduced into baculovirus using the "BaculoGold™" transfection kit (Pharmingen, San Diego, CA) using methods established by the manufacturer. Individual virus isolates are analyzed for protein production by radiolabeling infected cells with ³⁵S-methionine at 24 hours post infection. Infected cells are harvested at 48 hours post infection, and the labeled proteins are visualized by SDS-PAGE. Viruses exhibiting high expression levels can be isolated and used for scaled up expression.

For expression of a NgR polypeptide in a Sf9 cells, a polynucleotide molecule having the nucleotide sequence of SEQ ID NO:1 (or SEQ ID NO:3 or SEQ ID NO:13) can be amplified by PCR using the primers and methods described above for baculovirus expression. The NgR cDNA is cloned into vector pAcHLT-A (Pharmingen) for expression in Sf9 insect. The insert is cloned into the *NdeI* and *KpnI* sites, after elimination of an internal *NdeI* site (using the same primers described above for expression in baculovirus). DNA is purified with Qiagen chromatography columns and expressed in Sf9 cells. Preliminary Western blot experiments from non-purified plaques are tested for the presence of the recombinant protein of the expected size which reacted with the NgR-specific antibody. These results are confirmed after further purification and expression optimization in HiG5 cells.

F. Expression of soluble forms of NgR2 and NgR3 as NgR-Ig fusion proteins.

To generate a NgR2-Ig fusion protein, standard methods may be used as described in the literature (*e.g.* Sanicola *et al.* (1997) *Proc. Natl. Acad. Sci. USA.* 94, 6238-6243). For example, a DNA fragment encoding NgR2 without the sequence encoding the hydrophobic C-terminus (GPI anchor signal) may be ligated to a DNA fragment encoding the Fc domain of IgG1 (which may be human IgG1), and the chimeric fragment may be cloned into an expression vector to generate a plasmid. The plasmid may then be transfected into Chinese hamster ovary cells to generate a stable cell line producing the fusion protein. The fusion protein is then purified from conditioned media using standard methods. For example, clarified conditioned media from the cell line may be loaded by gravity directly onto Protein A Sepharose. The column may then be washed with five column volumes each of PBS, PBS containing

- 122 -

0.5 M NaCl, and 25 mM sodium phosphate, 100 mM NaCl (pH 5.0). The bound protein may then be eluted with 25 mM NaH₂PO₄, 100 mM NaCl (pH 2.8) and immediately neutralized with 1/10 fraction volume of 0.5 M Na₂HPO₄ (pH 8.6).

Similar methods may be used to generate a NgR3-Ig fusion protein.

5

Example 12: Interaction Trap/Two-Hybrid System

In order to assay for NgR-interacting proteins, the interaction trap/two-hybrid library screening method can be used. This assay was first described in Fields *et al.* (1989) *Nature* 340, 245, which is incorporated herein by reference in its entirety. A
10 protocol is published in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY 1999, John Wiley & Sons, NY and Ausubel, F. M. *et al.* 1992, SHORT PROTOCOLS IN MOLECULAR BIOLOGY, fourth edition, Greene and Wiley-interscience, NY, which is incorporated herein by reference in its entirety. Kits are available from Clontech, Palo Alto, CA (Matchmaker Two-Hybrid System 3).

15 A fusion of the nucleotide sequences encoding all or partial NgR and the yeast transcription factor GAL4 DNA-binding domain (DNA-BD) is constructed in an appropriate plasmid (*i.e.*, pGBKT7) using standard subcloning techniques. Similarly, a GAL4 active domain (AD) fusion library is constructed in a second plasmid (*i.e.*, pGADT7) from cDNA of potential NgR-binding proteins (for protocols on forming
20 cDNA libraries, see Sambrook *et al.* 1989, MOLECULAR CLONING: A LABORATORY MANUAL, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), which is incorporated herein by reference in its entirety. The DNA-BD/NgR fusion construct is verified by sequencing, and tested for autonomous reporter gene activation and cell toxicity, both of which would prevent a successful two-hybrid
25 analysis. Similar controls are performed with the AD/library fusion construct to ensure expression in host cells and lack of transcriptional activity. Yeast cells are transformed (ca. 105 transformants/mg DNA) with both the NgR and library fusion plasmids according to standard procedure (Ausubel, *et al.*, 1992, SHORT PROTOCOLS IN
MOLECULAR BIOLOGY, fourth edition, Greene and Wiley-interscience, NY, which is
30 incorporated herein by reference in its entirety). *In vivo* binding of DNA-BD/NgR with AD/library proteins results in transcription of specific yeast plasmid reporter genes (*i.e.*, *lacZ*, *HIS3*, *ADE2*, *LEU2*). Yeast cells are plated on nutrient-deficient

- 123 -

media to screen for expression of reporter genes. Colonies are dually assayed for β -galactosidase activity upon growth in Xgal (5-bromo-4-chloro-3-indolyl-b-D-galactoside) supplemented media (filter assay for β -galactosidase activity is described in Breeden *et al.*, (1985) *Cold Spring Harb. Symp. Quant. Biol.*, 50, 643, which is incorporated herein by reference in its entirety).

5 Positive AD-library plasmids are rescued from transformants and reintroduced into the original yeast strain as well as other strains containing unrelated DNA-BD fusion proteins to confirm specific NgR/library protein interactions. Insert DNA is sequenced to verify the presence of an open reading frame fused to GAL4 AD and to determine

10 the identity of the NgR-binding protein.

Example 13: Antibodies to Nogo-R

Standard techniques are employed to generate polyclonal or monoclonal antibodies to the NgR receptor, and to generate useful antigen-binding fragments

15 thereof or variants thereof, including "humanized" variants. Such protocols can be found, for example, in Sambrook *et al.* (1989), above, and Harlow *et al.* (Eds.), ANTIBODIES A LABORATORY MANUAL; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1988). In one embodiment, recombinant NgR polypeptides (or cells or cell membranes containing such polypeptides) are used as antigen to generate

20 the antibodies. In another embodiment, one or more peptides having amino acid sequences corresponding to an immunogenic portion of NgR (*e.g.*, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more amino acids) are used as antigen. Peptides corresponding to extracellular portions of Nogo-R, especially hydrophilic extracellular portions, are preferred. The antigen may be mixed with an adjuvant or linked to a

25 hapten to increase antibody production.

A. Polyclonal or Monoclonal antibodies

As one exemplary protocol, recombinant NgR or a synthetic fragment thereof is used to immunize a mouse for generation of monoclonal antibodies (or larger

30 mammal, such as a rabbit, for polyclonal antibodies). To increase antigenicity, peptides are conjugated to Keyhole Limpet Hemocyanin (Pierce), according to the manufacturer's recommendations. For an initial injection, the antigen is emulsified with

- 124 -

Freund's Complete Adjuvant and injected subcutaneously. At intervals of two to three weeks, additional aliquots of NgR antigen are emulsified with Freund's Incomplete Adjuvant and injected subcutaneously. Prior to the final booster injection, a serum sample is taken from the immunized mice and assayed by western blot to confirm the presence of antibodies that immunoreact with NgR. Serum from the immunized animals may be used as polyclonal antisera or used to isolate polyclonal antibodies that recognize NgR. Alternatively, the mice are sacrificed and their spleen removed for generation of monoclonal antibodies.

To generate monoclonal antibodies, the spleens are placed in 10 ml serum-free RPMI 1640, and single cell suspensions are formed by grinding the spleens in serum-free RPMI 1640, supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 100 units/ml penicillin, and 100 µg/ml streptomycin (RPMI) (Gibco, Canada). The cell suspensions are filtered and washed by centrifugation and resuspended in serum-free RPMI. Thymocytes taken from three naive Balb/c mice are prepared in a similar manner and used as a Feeder Layer. NS-1 myeloma cells, kept in log phase in RPMI with 10% fetal bovine serum (FBS) (Hyclone Laboratories, Inc., Logan, UT) for three days prior to fusion, are centrifuged and washed as well.

To produce hybridoma fusions, spleen cells from the immunized mice are combined with NS-1 cells and centrifuged, and the supernatant is aspirated. The cell pellet is dislodged by tapping the tube, and 2 ml of 37°C PEG 1500 (50% in 75 mM HEPES, pH 8.0) (Boehringer-Mannheim) is stirred into the pellet, followed by the addition of serum-free RPMI. Thereafter, the cells are centrifuged, resuspended in RPMI containing 15% FBS, 100 µM sodium hypoxanthine, 0.4 µM aminopterin, 16 µM thymidine (HAT) (Gibco), 25 units/ml IL-6 (Boehringer-Mannheim) and 1.5×10^6 thymocytes/ml, and plated into 10 Corning flat-bottom 96-well tissue culture plates (Corning, Corning, NY).

On days 2, 4, and 6 after the fusion, 100 µl of medium is removed from the wells of the fusion plates and replaced with fresh medium. On day 8, the fusions are screened by ELISA, testing for the presence of mouse IgG that binds to NgR. Selected fusion wells are further cloned by dilution until monoclonal cultures producing anti-NgR antibodies are obtained.

B. Humanization of anti-NgR monoclonal antibodies

The expression pattern of NgR as reported herein and the potential of NgRs as targets for therapeutic intervention suggest therapeutic indications for NgR inhibitors (antagonists). NgR-neutralizing antibodies comprise one class of therapeutics useful as
5 NgR antagonists. Following are protocols to improve the utility of anti-NgR monoclonal antibodies as therapeutics in humans by "humanizing" the monoclonal antibodies to improve their serum half-life and render them less immunogenic in human hosts (*i.e.*, to prevent human antibody response to non-human anti-NgR antibodies).

The principles of humanization have been described in the literature and are
10 facilitated by the modular arrangement of antibody proteins. To minimize the possibility of binding complement, a humanized antibody of the IgG4 isotype is preferred.

For example, a level of humanization is achieved by generating chimeric antibodies comprising the variable domains of non-human antibody proteins of interest
15 with the constant domains of human antibody molecules. (See, *e.g.*, Morrison *et al.*, (1989) *Adv. Immunol.*, 44, 65-92). The variable domains of NgR-neutralizing anti-NgR antibodies are cloned from the genomic DNA of a B-cell hybridoma or from cDNA generated from mRNA isolated from the hybridoma of interest. The V region gene fragments are linked to exons encoding human antibody constant domains, and
20 the resultant construct is expressed in suitable mammalian host cells (*e.g.*, myeloma or CHO cells).

To achieve an even greater level of humanization, only those portions of the variable region gene fragments that encode antigen-binding complementarity determining regions ("CDR") of the non-human monoclonal antibody genes are cloned
25 into human antibody sequences. (See, *e.g.*, Jones *et al.*, (1986) *Nature* 321, 522-525; Riechmann *et al.*, (1988) *Nature* 332, 323-327; Verhoeven *et al.*, (1988) *Science* 239, 1534-1536; and Tempest *et al.*, (1991) *Bio/Technology* 9, 266-271). If necessary, the β -sheet framework of the human antibody surrounding the CDR3 regions also is modified to more closely mirror the three dimensional structure of the antigen-binding
30 domain of the original monoclonal antibody. (See Kettleborough *et al.*, (1991) *Protein Engin.* 4, 773-783; and Foote *et al.*, (1992) *J. Mol. Biol.* 224, 487-499).

- 126 -

In an alternative approach, the surface of a non-human monoclonal antibody of interest is humanized by altering selected surface residues of the non-human antibody, *e.g.*, by site-directed mutagenesis, while retaining all of the interior and contacting residues of the non-human antibody. See Padlan (1991) *Mol. Immunol.* 28, 489-498 .

5 The foregoing approaches are employed using NgR-neutralizing anti-NgR monoclonal antibodies and the hybridomas that produce them to generate humanized NgR-neutralizing antibodies useful as therapeutics to treat or palliate conditions wherein NgR expression or ligand-mediated NgR signaling is detrimental.

10 **C. Human NgR-Neutralizing Antibodies from Phage Display**

Human NgR-neutralizing antibodies are generated by phage display techniques such as those described in Aujame *et al.* (1997) *Human Antibodies* 8, 155-168; Hoogenboom (1997) *TIBTECH* 15, 62-70; and Rader *et al.* (1997), *Curr. Opin. Biotechnol.* 8, 503-508, all of which are incorporated by reference. For example,
15 antibody variable regions in the form of Fab fragments or linked single chain Fv fragments are fused to the amino terminus of filamentous phage minor coat protein pIII. Expression of the fusion protein and incorporation thereof into the mature phage coat results in phage particles that present an antibody on their surface and contain the genetic material encoding the antibody. A phage library comprising such constructs is
20 expressed in bacteria, and the library is screened for NgR-specific phage-antibodies using labeled or immobilized NgR as antigen-probe.

D. Human NgR-neutralizing antibodies from transgenic mice

Human NgR-neutralizing antibodies are generated in transgenic mice essentially
25 as described in Bruggemann *et al.* (1996) *Immunol. Today* 17, 391-397 and Bruggemann *et al.* (1997) *Curr. Opin. Biotechnol.* 8, 455-458. Transgenic mice carrying human V-gene segments in germline configuration and that express these transgenes in their lymphoid tissue are immunized with a NgR composition using conventional immunization protocols. hybridomas are generated using B cells from the
30 immunized mice using conventional protocols and screened to identify hybridomas secreting anti-NgR human antibodies (*e.g.*, as described above).

Example 14: Assays to Identify Modulators of NgR Activity

Set forth below are several nonlimiting assays for identifying modulators (agonists and antagonists) of NgR activity. Among the modulators that can be identified by these assays are natural ligand compounds of the receptor; synthetic
5 analogs and derivatives of natural ligands; antibodies, antibody fragments, and/or antibody-like compounds derived from natural antibodies or from antibody-like combinatorial libraries; and/or synthetic compounds identified by high-throughput screening of libraries; and the like. All modulators that bind NgR are useful for identifying NgR in tissue samples (*e.g.*, for diagnostic purposes, pathological purposes,
10 and the like). Agonist and antagonist modulators are useful for up-regulating and down-regulating NgR activity, respectively, to treat disease states characterized by abnormal levels of NgR activity. The assays may be performed using single putative modulators, and/or may be performed using a known agonist in combination with candidate antagonists (or *visa versa*).

15

A. cAMP Assays

In one type of assay, levels of cyclic adenosine monophosphate (cAMP) are measured in NgR-transfected cells that have been exposed to candidate modulator compounds. Protocols for cAMP assays have been described in the literature. (See,
20 *e.g.*, Sutherland *et al.*, (1968) *Circulation* 37, 279; Frandsen *et al.*, (1976) *Life Sciences* 18, 529-541; Dooley *et al.*, (1997) *J. Pharmacol. Exp. Therap.* 283, 735-41; and George *et al.*, (1997) *J. Biomol. Screening* 2, 235-40). An exemplary protocol for such an assay, using an Adenylyl Cyclase Activation FlashPlate® Assay from NEN™ Life Science Products, is set forth below.

25 Briefly, the NgR coding sequence (*e.g.*, a cDNA or intronless genomic DNA) is subcloned into a commercial expression vector, such as pzeoSV2 (Invitrogen), and transiently transfected into Chinese Hamster Ovary (CHO) cells using known methods, such as the transfection protocol provided by Boehringer-Mannheim when supplying the FuGENE 6 transfection reagent. Transfected CHO cells are seeded into 96-well
30 microplates from the FlashPlate® assay kit, which are coated with solid scintillant to which antisera to cAMP has been bound. For a control, some wells are seeded with

- 128 -

wild type (untransfected) CHO cells. Other wells in the plate receive various amounts of a cAMP standard solution for use in creating a standard curve.

One or more test compounds (*i.e.*, candidate modulators) are added to the cells in each well, with water and/or compound-free medium/diluent serving as a control or
5 controls. After treatment, cAMP is allowed to accumulate in the cells for exactly 15 minutes at room temperature. The assay is terminated by the addition of lysis buffer containing [¹²⁵I]-labeled cAMP, and the plate is counted using a Packard Topcount™ 96-well microplate scintillation counter. Unlabeled cAMP from the lysed cells (or
10 from standards) and fixed amounts of [¹²⁵I]-cAMP compete for antibody bound to the plate. A standard curve is constructed, and cAMP values for the unknowns are obtained by interpolation. Changes in intracellular cAMP levels of cells in response to exposure to a test compound are indicative of NgR modulating activity. Modulators that act as agonists of receptors which couple to the G_s subtype of G proteins will stimulate production of cAMP, leading to a measurable 3-10 fold increase in cAMP
15 levels. Agonists of receptors which couple to the G_{i/o} subtype of G proteins will inhibit forskolin-stimulated cAMP production, leading to a measurable decrease in cAMP levels of 50-100%. Modulators that act as inverse agonists will reverse these effects at receptors that are either constitutively active or activated by known agonists.

20 B. Aequorin Assays

In another assay, cells (*e.g.*, CHO cells) are transiently co-transfected with both a NgR expression construct and a construct that encodes the photoprotein apoaequorin. In the presence of the cofactor coelenterazine, apoaequorin will emit a measurable luminescence that is proportional to the amount of intracellular (cytoplasmic) free
25 calcium. (See generally, Cobbold, *et al.* "Aequorin measurements of cytoplasmic free calcium," *In*: McCormack J.G. and Cobbold P.H., eds., CELLULAR CALCIUM: A PRACTICAL APPROACH. Oxford:IRL Press (1991); Stables *et al.*, (1997) *Anal. Biochem.* 252, 115-26; and Haugland, HANDBOOK OF FLUORESCENT PROBES AND RESEARCH CHEMICALS. Sixth edition. Molecular Probes, Eugene, OR (1996)).

30 In one exemplary assay, NgR is subcloned into the commercial expression vector pzeoSV2 (Invitrogen) and transiently co-transfected along with a construct that encodes the photoprotein apoaequorin (Molecular Probes, Eugene, OR) into CHO cells

- 129 -

using the transfection reagent FuGENE 6 (Boehringer-Mannheim) and the transfection protocol provided in the product insert.

The cells are cultured for 24 hours at 37°C in MEM (Gibco/BRL, Gaithersburg, MD) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin, at which time the medium is changed to serum-free MEM containing 5 µM coelenterazine (Molecular Probes, Eugene, OR). Culturing is then continued for two additional hours at 37°C. Subsequently, cells are detached from the plate using VERSEN (Gibco/BRL), washed, and resuspended at 200,000 cells/ml in serum-free MEM.

Dilutions of candidate NgR modulator compounds are prepared in serum-free MEM and dispensed into wells of an opaque 96-well assay plate at 50 µl/well. Plates are then loaded onto an MLX microtiter plate luminometer (Dynex Technologies, Inc., Chantilly, VA). The instrument is programmed to dispense 50 µl cell suspensions into each well, one well at a time, and immediately read luminescence for 15 seconds.

Dose-response curves for the candidate modulators are constructed using the area under the curve for each light signal peak. Data are analyzed with SlideWrite, using the equation for a one-site ligand, and EC₅₀ values are obtained. Changes in luminescence caused by the compounds are considered indicative of modulatory activity. Modulators that act as agonists at receptors which couple to the G_q subtype of G proteins give an increase in luminescence of up to 100 fold. Modulators that act as inverse agonists will reverse this effect at receptors that are either constitutively active or activated by known agonists.

C. Luciferase Reporter Gene Assay

The photoprotein luciferase provides another useful tool for assaying for modulators of NgR activity. Cells (*e.g.*, CHO cells or COS 7 cells) are transiently co-transfected with both a NgR expression construct (*e.g.*, NgR in pzeoSV2) and a reporter construct which includes a gene for the luciferase protein downstream from a transcription factor binding site, such as the cAMP-response element (CRE), AP-1, or NF-kappa B. Expression levels of luciferase reflect the activation status of the signaling events. (See generally, George *et al.* (1997) *J. Biomol. Screening* 2, 235-240; and Stratowa *et al.* (1995) *Curr. Opin. Biotechnol.* 6, 574-581). Luciferase

- 130 -

activity may be quantitatively measured using, *e.g.*, luciferase assay reagents that are commercially available from Promega (Madison, WI).

In one exemplary assay, CHO cells are plated in 24-well culture dishes at a density of 100,000 cells/well one day prior to transfection and cultured at 37°C in
5 MEM (Gibco/BRL) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin. Cells are transiently co-transfected with both a NgR expression construct and a reporter construct containing the luciferase gene. The reporter plasmids CRE-luciferase, AP-1-luciferase and NF-kappaB-luciferase may be purchased from Stratagene (Legally, CA). Transfections are
10 performed using the FuGENE 6 transfection reagent (Boehringer-Mannheim) according to the supplier's instructions. Cells transfected with the reporter construct alone are used as a control. Twenty-four hours after transfection, cells are washed once with PBS pre-warmed to 37°C. Serum-free MEM is then added to the cells either alone (control) or with one or more candidate modulators and the cells are
15 incubated at 37°C for five hours. Thereafter, cells are washed once with ice-cold PBS and lysed by the addition of 100 µl of lysis buffer per well from the luciferase assay kit supplied by Promega. After incubation for 15 minutes at room temperature, 15 µl of the lysate is mixed with 50 µl of substrate solution (Promega) in an opaque-white, 96-well plate, and the luminescence is read immediately on a Wallace model 1450
20 MicroBeta scintillation and luminescence counter (Wallace Instruments, Gaithersburg, MD).

Differences in luminescence in the presence versus the absence of a candidate modulator compound are indicative of modulatory activity. Receptors that are either constitutively active or activated by agonists typically give a 3-20-fold stimulation of
25 luminescence compared to cells transfected with the reporter gene alone. Modulators that act as inverse agonists will reverse this effect.

D. Intracellular calcium measurement using FLIPR

Changes in intracellular calcium levels are another recognized indicator of
30 receptor activity, and such assays can be employed to screen for modulators of NgR activity. For example, CHO cells stably transfected with a NgR expression vector are plated at a density of 4×10^4 cells/well in Packard black-walled, 96-well plates

- 131 -

5 specially designed to discriminate fluorescence signals emanating from the various wells on the plate. The cells are incubated for 60 minutes at 37°C in modified Dulbecco's PBS (D-PBS) containing 36 mg/L pyruvate and 1 g/L glucose with the addition of 1% fetal bovine serum and one of four calcium indicator dyes (Fluo-3™ AM, Fluo-4™ AM, Calcium Green™ -1 AM, or Oregon Green™ 488 BAPTA-1 AM), each at a concentration of 4 μM. Plates are washed once with modified D-PBS without 1% fetal bovine serum and incubated for 10 minutes at 37°C to remove residual dye from the cellular membrane. In addition, a series of washes with modified D-PBS without 1% fetal bovine serum is performed immediately prior to activation of the calcium response.

A calcium response is initiated by the addition of one or more candidate receptor agonist compounds, calcium ionophore A23187 (10 μM; positive control), or ATP (4 μM; positive control). Fluorescence is measured by Molecular Device's FLIPR with an argon laser (excitation at 488 nm). (See, *e.g.*, Kuntzweiler *et al.* (1998) *Drug Dev. Res.* 44,14-20). The F-stop for the detector camera is set at 2.5 and the length of exposure is 0.4 milliseconds. Basal fluorescence of cells is measured for 20 seconds prior to addition of candidate agonist, ATP, or A23187, and the basal fluorescence level is subtracted from the response signal. The calcium signal is measured for approximately 200 seconds, taking readings every two seconds. Calcium ionophore A23187 and ATP increase the calcium signal 200% above baseline levels. In general, activated NgRs increase the calcium signal at least about 10-15% above baseline signal.

E. [³⁵S]GTPγS Binding Assay

25 It is also possible to evaluate whether NgR signals through a G protein-mediated pathway. Because G protein-coupled receptors signal through intracellular G proteins whose activity involves GTP binding and hydrolysis to yield bound GDP, measurement of binding of the non-hydrolyzable GTP analog [³⁵S]-GTPγS in the presence and absence of candidate modulators provides another assay for modulator activity. (See, *e.g.*, Kowal *et al.*, (1998) *Neuropharmacology* 37, 179-187.).

- 132 -

In one exemplary assay, cells stably transfected with a NgR expression vector are grown in 10 cm tissue culture dishes to subconfluence, rinsed once with 5 ml of ice-cold $\text{Ca}^{2+}/\text{Mg}^{2+}$ -free phosphate-buffered saline, and scraped into 5 ml of the same buffer. Cells are pelleted by centrifugation (500 x g, 5 minutes), resuspended in TEE
5 buffer (25 mM Tris, pH 7.5, 5 mM EDTA, 5 mM EGTA), and frozen in liquid nitrogen. After thawing, the cells are homogenized using a Dounce homogenizer (1 ml TEE per plate of cells), and centrifuged at 1,000 x g for 5 minutes to remove nuclei and unbroken cells.

The homogenate supernatant is centrifuged at 20,000 x g for 20 minutes to
10 isolate the membrane fraction, and the membrane pellet is washed once with TEE and resuspended in binding buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM MgCl_2 , 1 mM EDTA). The resuspended membranes can be frozen in liquid nitrogen and stored at -70°C until use.

Aliquots of cell membranes prepared as described above and stored at -70°C
15 are thawed, homogenized, and diluted into buffer containing 20 mM HEPES, 10 mM MgCl_2 , 1 mM EDTA, 120 mM NaCl, 10 μM GDP, and 0.2 mM ascorbate, at a concentration of 10-50 $\mu\text{g}/\text{ml}$. In a final volume of 90 μl , homogenates are incubated with varying concentrations of candidate modulator compounds or 100 μM GTP for 30 minutes at 30°C and then placed on ice. To each sample, 10 μl guanosine
20 $5'$ -O-(3 ^{35}S)thio) triphosphate (NEN, 1200 Ci/mmol; [^{35}S]-GTP γS), was added to a final concentration of 100-200 pM. Samples are incubated at 30°C for an additional 30 minutes, 1 ml of 10 mM HEPES, pH 7.4, 10 mM MgCl_2 , at 4°C is added and the reaction is stopped by filtration.

Samples are filtered over Whatman GF/B filters and the filters are washed with
25 20 ml ice-cold 10 mM HEPES, pH 7.4, 10 mM MgCl_2 . Filters are counted by liquid scintillation spectroscopy. Nonspecific binding of [^{35}S]-GTP γS is measured in the presence of 100 μM GTP and subtracted from the total. Compounds are selected that modulate the amount of [^{35}S]-GTP γS binding in the cells, compared to untransfected control cells. Activation of receptors by agonists gives up to a five-fold increase in
30 [^{35}S]-GTP γS binding. This response is blocked by antagonists.

F. [^3H]Arachidonic Acid Release

- 133 -

The activation of NgRs may also potentiate arachidonic acid release in cells, providing yet another useful assay for modulators of NgR activity. (See, e.g., Kanterman *et al.*, (1991) *Mol. Pharmacol.* 39,364-369.) For example, CHO cells that are stably transfected with a NgR expression vector are plated in 24-well plates at a density of 15,000 cells/well and grown in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin for 48 hours at 37°C before use. Cells of each well are labeled by incubation with [³H]-arachidonic acid (Amersham Corp., 210 Ci/mmol) at 0.5 µCi/ml in 1 ml MEM supplemented with 10 mM HEPES, pH 7.5, and 0.5% fatty-acid-free bovine serum albumin for 2 hours at 37°C. The cells are then washed twice with 1 ml of the same buffer.

Candidate modulator compounds are added in 1 ml of the same buffer, either alone or with 10 µM ATP and the cells are incubated at 37°C for 30 minutes. Buffer alone and mock-transfected cells are used as controls. Samples (0.5 ml) from each well are counted by liquid scintillation spectroscopy. Agonists which activate the receptor will lead to potentiation of the ATP-stimulated release of [³H]-arachidonic acid. This potentiation is blocked by antagonists.

G. Extracellular Acidification Rate

In yet another assay, the effects of candidate modulators of NgR activity are assayed by monitoring extracellular changes in pH induced by the test compounds (see, e.g., Dunlop *et al.* (1998) *J. Pharmacol. Toxicol. Meth.* 40, 47-55). In one embodiment, CHO cells transfected with a NgR expression vector are seeded into 12 mm capsule cups (Molecular Devices Corp.) at 4 x 10⁵ cells/cup in MEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 10 U/ml penicillin, and 10 µg/ml streptomycin. The cells are incubated in this medium at 37°C in 5% CO₂ for 24 hours.

Extracellular acidification rates are measured using a Cytosensor microphysiometer (Molecular Devices Corp.). The capsule cups are loaded into the sensor chambers of the microphysiometer and the chambers are perfused with running buffer (bicarbonate-free MEM supplemented with 4 mM L-glutamine, 10 units/ml penicillin, 10 µg/ml streptomycin, 26 mM NaCl) at a flow rate of 100 µl/minute.

- 134 -

Candidate agonists or other agents are diluted into the running buffer and perfused through a second fluid path. During each 60-second pump cycle, the pump is run for 38 seconds and is off for the remaining 22 seconds. The pH of the running buffer in the sensor chamber is recorded during the cycle from 43-58 seconds, and the pump is re-started at 60 seconds to start the next cycle. The rate of acidification of the running buffer during the recording time is calculated by the Cytosoft program. Changes in the rate of acidification are calculated by subtracting the baseline value (the average of 4 rate measurements immediately before addition of a modulator candidate) from the highest rate measurement obtained after addition of a modulator candidate. The selected instrument detects 61 mV/pH unit. Modulators that act as agonists of the receptor result in an increase in the rate of extracellular acidification compared to the rate in the absence of agonist. This response is blocked by modulators which act as antagonists of the receptor.

Example 15: mNgR3 does not bind hNogo-A(1055-1120)

To functionally test the mouse NgR3 (hereinafter, mNgR3) for its ability to bind hNogo-A(1055-1120), a cDNA expression vector for a myc epitope-tagged mNgR3 protein was created. The mouse NgR3 cDNA was amplified by PCR from mouse adult brain cDNA, from the signal sequence to the stop codon, and ligated into the pSecTag2 vector such that the vector encodes a signal sequence followed by a myc tag followed by the mature mNgR3 sequence. This plasmid was transfected into COS7 cells, and expression of a myc-tagged protein of the predicted size was verified by immunoblot analysis. Alkaline phosphatase-hNogo-A(1055-1120) binding studies and myc immunohistology were conducted as described (Fournier et al., supra).

The cells expressing mNgR3 express the myc-tagged protein but binding to AP-hNogo-A(1055-1120) was not observed under the conditions employed (Fig. 8).

Example 16: Identification of partial human NgR3 cDNA and protein sequences

The tblastn program was used to search for the human homolog of mouse NgR3. The mouse NgR3 protein sequence (SEQ ID NO:4) was used to query a proprietary human expressed sequence tag (EST) database from Incyte yielding one highly significant hit: Incyte Template ID 190989.1. This sequence (937 nucleotides)

- 135 -

contains an open reading frame of 312 amino acids in the second reverse frame that exhibits 88% identity with residues 66 to 381 of mouse NgR3 (SEQ ID NO:4), strongly indicating that it is part of the human NgR3 homolog.

A query of SEQ ID NO:4 against the public human EST database in Genbank
5 also produced a hit with a 465-bp EST (Accession number: R35699; Version number: R35699.1; GI: 792600). There are a number of single nucleotide deletions and insertions within this sequence which cause frame shift errors. All of the reliable sequence contained in this public EST is present in the Incyte EST (Template ID 190989.1).

10 To obtain more nucleotide sequence that would extend the amino acid sequence at that carboxy terminal end, the I.M.A.G.E. Consortium clone No. 38319, which corresponds to Genbank accession No. R35699, was purchased from Incyte Genomics Inc. and subjected to further DNA sequence analysis. This clone consists of a NotI/HinD III fragment containing the sequence of interest, cloned into the
15 NotI/HinD III sites of the vector Lafmid BA (<http://image.llnl.gov/image/html/libs/lafmidBA.shtml>). The clone was received as an agar stab, which was streaked out on LB agar plates containing 50ug/ml ampicillin to isolate individual colonies. Six colonies were grown in LB medium with antibiotic, and plasmid DNA was prepared using the Promega Wizard Plus Miniprep DNA
20 Purification System (Promega #A7500). These DNAs were subsequently digested with NotI and HinD III restriction enzymes to confirm that the clones contained an insert. The insert of one isolate was sequenced using a combination of vector specific and gene specific primers yielding a partial nucleotide sequence of human NgR3 of 1176 nucleotides (SEQ ID NO:13). A translation of this sequence provides a partial
25 sequence for human NgR3 of 392 amino acids (SEQ ID NO:14).

The nucleotide sequence of SEQ ID NO:13 differs from the Incyte EST sequence at three positions. Nucleotide positions 12-13 in SEQ ID NO:13 are CG, whereas the corresponding nucleotides in the Incyte Template ID 190989.1 are GT (i.e., positions 12-13 of the complement of Incyte Template ID 190989.1). In
30 addition, position 641 in SEQ ID NO:13 is a C, whereas the corresponding nucleotide in the Incyte Template ID 190989.1 sequence is an A (i.e., position 641 of the complement of Incyte Template ID 190989.1). This results in two changes in amino

- 136 -

acids when comparing SEQ ID NO:14 to the ORF encoded by Incyte Template 190989.1: SEQ ID NO:14 contains a valine at position 5, whereas the ORF encoded by Incyte Template ID 190989.1 contains a leucine; SEQ ID NO:14 contains an alanine at position 214, whereas the ORF encoded by Incyte Template ID 190989.1 contains a glutamic acid.

The nucleotide sequence of SEQ ID NO:13 differs from the public EST (Accession number: R35699; Version number: R35699.1; GI: 792600) sequence at two positions (within the first 200 nucleotides of reliable sequence). Nucleotide positions 12-13 in SEQ ID NO:13 are CG, whereas the corresponding nucleotides in the public EST are GT (i.e., positions 12-13 of the public EST; Accession no: R35699; Version no: R35699.1; GI: 792600) This leads to a single amino acid change when comparing SEQ ID NO:14 to the ORF encoded by the public EST: SEQ ID NO:14 contains a valine at position 5, while the ORF encoded by the public EST contains a leucine.

A Bestfit analysis of the partial human amino acid sequence with the full-length mouse amino acid sequence indicates that the human NgR3 amino acid sequence is complete at the carboxy terminal end and that they share 89.54% identity. An alignment of all the NgR proteins is shown in Figure 9. Although the human NgR3 amino acid sequence is missing the first 25 amino acids, it can be determined that the human NgR3 protein contains the following features in common with the other NgR sequences: (1) eight Leucine Rich Repeat (LRR) domains; (2) an LRR carboxy-terminal (LRR-CT) domain; (3) a conserved cysteine in the fourth LRR domain; (4) a conserved potential glycosylation site in the eighth LRR domain; and (5) a hydrophobic carboxyl terminus.

As those skilled in the art will appreciate, numerous changes and modifications may be made to the preferred embodiments of the invention without departing from the spirit of the invention. It is intended that all such variations fall within the scope of the invention.

The entire disclosure of each publication cited herein is hereby incorporated by reference. This application claims benefit from United States provisional application 60/238,361, filed October 6, 2000, which is incorporated by reference herein in its entirety.

- 137 -

Key for Sequence Listing:

	SEQ ID NO:1	human NgR2 cDNA sequence derived from genomic sequence AC013606
	SEQ ID NO:2	human NgR2 amino acid sequence
5	SEQ ID NO:3	mouse NgR3 cDNA sequence derived from AC021768
	SEQ ID NO:4	a mouse NgR3 amino acid sequence
	SEQ ID NO:5	a human NgR1 amino acid sequence
	SEQ ID NO:6	a consensus amino acid sequence for NgRs
	SEQ ID NO:7	#1055-1120 amino acid residues of hNogoA (Nogo-66)
10	SEQ ID NO:8	a mature human NgR2 amino acid sequence
	SEQ ID NO:9	a mature mouse NgR3 amino acid sequence
	SEQ ID NO:10	a consensus NgR LLRNT amino acid sequence
	SEQ ID NO:11	a consensus NgR LRRCT domain amino acid sequence
	SEQ ID NO:12	a consensus NgR LRR domain amino acid sequence
15	SEQ ID NO:13	a partial human NgR3 nucleotide sequence
	SEQ ID NO:14	a partial human NgR3 amino acid sequence
	SEQ ID NO:15	a genomic sequence encoding a human NgR2 sequence.
	SEQ ID NO:16	a genomic sequence (complementary strand) encoding a mouse NgR3
20	SEQ ID NO:17	a mouse NgR1 amino acid sequence
	SEQ ID NO:18	a consensus sequence for the NTLRRCT domain of NgR
	SEQ ID NO:19	an consensus NgR LRRCT domain amino acid sequence

25

30

- 138 -

CLAIMS

What is claimed is:

5

1. An isolated nucleic acid comprising a nucleotide sequence encoding a polypeptide comprising an LRRCT domain consisting of the amino acid sequence:

10

N X₁ W X₂ C X₃ C R A R X₄ L W X₅ W X₆ X₇ X₈ X₉ R X₁₀ S S S X₁₁ V

X₁₂ C X₁₃ X₁₄ P X₁₅ X₁₆ X₁₇ X₁₈ X₁₉ X₂₀ D L X₂₁ X₂₂ L X₂₃ X₂₄ X₂₅ D

X₂₆ X₂₇ X₂₈ C [SEQ ID NO: 19]

15

wherein X is any amino acid or a gap and the polypeptide does not comprise the amino acid sequence from residue 260 to 309 of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1).

20

2. The isolated nucleic acid according to claim 1, wherein X₁₇ and X₂₃ are independently selected from the group consisting of: arginine and lysine.

25

3. The isolated nucleic acid according to claim 2, wherein the amino acid sequence of the LRRCT domain is selected from the group consisting of: residues #261-310 of SEQ ID NO:2 and residues 261-310 of SEQ ID NO: 2 with up to 10 conservative amino acid substitutions.

30

4. An isolated nucleic acid encoding the polypeptide of SEQ ID NO: 2.

5. An isolated nucleic acid encoding the polypeptide of SEQ ID NO: 4 (mouse NgR3) or SEQ ID NO: 14 (human NgR3).

6. The isolated nucleic acid according to claim 1, wherein the

- 139 -

polypeptide comprises: (a) a NTLRRCT domain, and (b) less than a complete CTS domain, provided that a partial CTS domain, if present, consists of no more than the first 39 amino acids of the CTS domain.

5 7. The isolated nucleic acid to claim 1, wherein the polypeptide does not comprise an intact GPI domain.

8. An isolated nucleic acid consisting essentially of a nucleotide sequence complementary to a nucleotide sequence encoding a polypeptide selected
10 from the group consisting of: a polypeptide consisting of residues 311-395 of SEQ ID NO: 2, a polypeptide consisting of residues 256-396 of SEQ ID NO: 14 and a polypeptide consisting of residues 321-438 of SEQ ID NO: 4, wherein the nucleic acid is from 8 to 100 nucleotides in length.

15 9. A vector comprising the nucleic acid of any one of claims 1, 4 or 5.

10. A host cell comprising a vector according to claim 9.

11. A polypeptide comprising a LRRCT amino acid sequence:

20

N X₁ W X₂ C X₃ C R A R X₄ L W X₅ W X₆ X₇ X₈ X₉ R X₁₀ S S S X₁₁ V

X₁₂ C X₁₃ X₁₄ P X₁₅ X₁₆ X₁₇ X₁₈ X₁₉ X₂₀ D L X₂₁ X₂₂ L X₂₃ X₂₄ X₂₅ D

25 X₂₆ X₂₇ X₂₈ C [SEQ ID NO: 19]

wherein X is any amino acid residue or a gap and the polypeptide does not comprise the amino acid sequence from residue 260 to 309 of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1).

30

12. The polypeptide according to claim 11, wherein X₁₇ and X₂₃ is selected from the group consisting of arginine and lysine.

- 140 -

13. The polypeptide according to claim 11, wherein X₁₉ is glycine.

[SEQ ID NO:11]

14. The polypeptide according to claim 11, wherein the amino acid
5 sequence is selected from the group consisting of residues 261–310 of SEQ ID NO:2, residues 206–255 of SEQ ID NO: 14, residues 271–320 of SEQ ID NO:4 and amino acid sequences thereof comprising a conservative substitution.

15. A polypeptide comprising a NTLRRCT amino acid sequence:

10

CPX₁X₂CX₃CYX₄X₅PX₆X₇TX₈SCX₉X₁₀X₁₁X₁₂X₁₃X₁₄X₁₅X₁₆P
X₁₇X₁₈X₁₉PX₂₀X₂₁X₂₂X₂₃RX₂₄FLX₂₅X₂₆NX₂₇IX₂₈X₂₉X₃₀X₃₁X₃₂X₃₃
X₃₄FX₃₅X₃₆X₃₇X₃₈X₃₉X₄₀X₄₁X₄₂LWX₄₃X₄₄SNX₄₅X₄₆X₄₇X₄₈IX₄₉
X₅₀X₅₁X₅₂FX₅₃X₅₄X₅₅X₅₆X₅₇LEX₅₈LDLX₅₉DNX₆₀X₆₁LX₆₂X₆₃X₆₄
15 X₆₅PX₆₆TFX₆₇GLX₆₈X₆₉LX₇₀X₇₁LX₇₂LX₇₃X₇₄CX₇₅LX₇₆X₇₇LX₇₈
X₇₉X₈₀X₈₁FX₈₂GLX₈₃X₈₄LQYLYLQX₈₅NX₈₆X₈₇X₈₈X₈₉LX₉₀D
X₉₁X₉₂FX₉₃DLX₉₄NLX₉₅HLFLHGNX₉₆X₉₇X₉₈X₉₉X₁₀₀X₁₀₁X₁₀₂
X₁₀₃X₁₀₄FRGLX₁₀₅X₁₀₆LDRLLLHX₁₀₇NX₁₀₈X₁₀₉X₁₁₀X₁₁₁VHX₁₁₂
X₁₁₃AFX₁₁₄X₁₁₅LX₁₁₆RLX₁₁₇X₁₁₈LX₁₁₉LFX₁₂₀NX₁₂₁LX₁₂₂X₁₂₃L
20 X₁₂₄X₁₂₅X₁₂₆X₁₂₇LX₁₂₈X₁₂₉LX₁₃₀X₁₃₁LX₁₃₂X₁₃₃LRLNX₁₃₄NX₁₃₅W
X₁₃₆CX₁₃₇CRX₁₃₈RX₁₃₉LWX₁₄₀WX₁₄₁X₁₄₂X₁₄₃X₁₄₄RX₁₄₅SSSX₁₄₆
VX₁₄₇CX₁₄₈X₁₄₉PX₁₅₀X₁₅₁X₁₅₂X₁₅₃X₁₅₄X₁₅₅DLX₁₅₆X₁₅₇LX₁₅₈X₁₅₉X₁₆₀
DX₁₆₁X₁₆₂X₁₆₃C [SEQ ID NO:18]

20

25

wherein X is any amino acid residue or a gap and wherein the polypeptide is not the polypeptide of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1).

30

16. The polypeptide according to claim 15, wherein X₆, X₃₇ and X₃₈ represents a gap.

17. A polypeptide comprising an amino sequence selected from the

- 141 -

group consisting of: SEQ ID NO:2, SEQ ID NO:4 and SEQ ID NO:14.

18. The polypeptide according any one of claims 11, 15 or 17, wherein the polypeptide comprises: (a) an NTLRRCT domain, and (b) less than a complete
5 CTS domain, provided that a partial CTS domain, if present, consists of no more than the first 39 amino acids of the CTS domain.

19. The polypeptide according to any one of claims 11, 15 or 17, wherein the polypeptide does not comprise an intact GPI domain.

10

20. The polypeptide according to any one of claims 11, 15 or 17, wherein the amino acid sequence of the polypeptide further comprises an amino acid sequence of a heterologous polypeptide.

21. The polypeptide according to claim 20, wherein the heterologous polypeptide is an Fc portion of an antibody.

22. A method of producing a polypeptide according to any one of claims 11, 15 or 17, comprising the steps of introducing an isolated nucleic acid
20 according to any one of claims 1, 4, 5 or 8 or a vector according to claim 9 into a host cell, culturing said host cell under conditions suitable for expression of said polypeptide, and recovering said polypeptide.

23. An antibody that binds to a polypeptide of any one of claims 11, 15
25 or 17.

24. A composition comprising the polypeptide of claim 11, 15 or 17 and a pharmaceutically acceptable carrier.

25. A composition comprising the antibody of claim 23 and a pharmaceutically acceptable carrier.

30

- 142 -

26. A method of decreasing inhibition of axonal growth of a CNS neuron, comprising the step of contacting the neuron with an effective amount of the polypeptide of claim 11, 15 or 17.

5 27. A method of treating a central nervous system disease, disorder or injury, comprising administering to a mammal an effective amount of the polypeptide of claim 11, 15 or 17.

10 28. A method of decreasing inhibition of axonal growth of a CNS neuron comprising the step of contacting the neuron with an effective amount of the antibody according to claim 23.

15 29. A method of treating a central nervous system disease, disorder or injury, comprising administering to a mammal an effective amount of the antibody according to claim 23.

30. A method for identifying a molecule that binds a polypeptide of claim 11, 15 or 17 comprising the steps of:

- 20 (a) providing a polypeptide of claim 11, 15 or 17;
 (b) contacting the polypeptide with the candidate molecule;
 and
 (c) detecting binding of the candidate molecule to the polypeptide.

25

FIG. 1A

FIG. 1A

NOCO-R2	1	~~~~~MLPG	LRRLLQAPAS	AC...LLLML	LA..LPLAAP	SCPMLCTCYS	50
NOCO-R3		MSWQSGTTVT	QSPVQAAQVS	GCCVELLLLL	LAGEPLGG	CCPRDCVCYP	
NOCO-R1		~~~~~MKRAS	AGSRLLAWV	LWLQAWQVAA	PCPGACVCYN		
Consensus		~~~~~S	-----S	-----LL---	L-----	-CP--C-CY-	
NOCO-R2	51	SP.PTVSCQA	NNFSSVPLSL	PPSTQRLFLQ	NNLIRTLRPG	TFGS...NLLT	100
NOCO-R3		AP.MTVSCQA	HNFAAIPEGI	PEDSERIFLQ	NNRITFLQQG	HFSP...AMVT	
NOCO-R1		EPKVTTSCPQ	QGLQAVPVGI	PAASQRIFLH	GNRISHVPAA	SFRACRNLT	
Consensus		-P--T-SC--	-----P---	P-----R-FL-	N-I-----	-F-----	
NOCO-R2	101	LWLFSSNNLST	IYPGTFRHLQ	ALEELDGLDN	RHLRSLEPDT	FQGLERLQSL	150
NOCO-R3		LWIYSNNITF	IAPNTFEGFV	HLEELDGLDN	RQLRTLAPET	FQGLVKLHAL	
NOCO-R1		LWLHSNVLAR	IDAAAFGLA	LLEQLDLSN	AQLRSVDPAT	FHGLGRLHTL	
Consensus		LW--SN----	I-----F----	-LE-LDL-DN	--LR---P-T	F-GL--L--L	
NOCO-R2	151	HLYRCQLSSL	PGNIFRGLVS	LQYLYQENS	LLHLQDDLFA	DLANLSHLFL	200
NOCO-R3		YLYKCGLSAL	PAGIFGGLHS	LQYLYQDNH	IEYLQDDIFV	DLVNLSHLFL	
NOCO-R1		HLDRCGLQEL	GPGLFRGLAA	LQYLYQDNA	LQALPDDTFR	DLGNLTHLFL	
Consensus		-L--C-L--L	-----F-GL--	LQYLYQ-N-	---L-DD-F-	DL-NL-HLFL	
NOCO-R2	201	HGNRLRLLTE	HVFRGLGSLD	RLLHGNRLQ	GVHRAAFRGL	SRLTILYLFN	250
NOCO-R3		HGNKLWSLQ	GIFRGLVNL	RLLHENQLQ	WVHKAFHDL	HRLTTLFLFN	
NOCO-R1		HGNRISSVPE	RAFRGLHSLD	RLLHQNRVA	HVHPAFRDL	GRMLTYLFA	
Consensus		HGN-----	--FRGL--LD	RLLH-N---	-VH--AF--L	-RL--L-LF-	

FIG. 1B

Fig. 1B

NOC0-R2	251	NSLASLPGEA	LADLPSLEFL	RLNANPWACD	CRARPLWAWF	QRARVSSSDV	300
NOC0-R3		NSLTELQDCD	LAPLVALEFL	RLNGNAWDCG	CRARSLWEWL	RRFRGSSSAV	
NOC0-R1		NNLSALPTEA	LAPLRALQYL	RLNDNPWVCD	CRARPLWAWL	QKFRGSSSEV	
Consensus		N-L--L----	LA-L--L--L	RLN-N-W-C-	CRAR-LW-W-	--R-SSS-V	
NOC0-R2	301	TCATPPERQG	RDLRALREAD	FQAC...P.P	AAPTRPGSRA	350
NOC0-R3		PCATPELRQG	QDLKLLRVED	FRNC...TGP	VSPHQIKSHT	
NOC0-R1		PCSLPQRLAG	RDLKRLAAND	LQCAVATGP	YHPIWTGRAT	DEEPLGLPKC	
Consensus		-C--P-----G	-DL--L---D	---C-----P	--P-----	-----	
NOC0-R2	351RGN	..SSNH.LY	G.VAE.....	AGAPPADPS.	..TLYRDLPA	400
NOC0-R3		..LTTSDDRAA	..RKEHHP SH	G.ASRDKGHP	HGHPPGSRSG	YKKAGKNCTS	
NOC0-R1		CQPDAAADKAS	VLEPGRPASA	GNALKGRVPP	GDSPPGNGSG	PRHI.NDSPF	
Consensus		-----	-----G	-----	--PP---S-	-----	
NOC0-R2	401	EDSRGR.....	QGGDApte.D	DYWGGY.....GGED	QRGEQMCPCA	450
NOC0-R3		HRNRNQISKV	SSGKELTELQ	DYAPDYQHKF	SFDIMPTARP	KRKGKCART	
NOC0-R1		GTLPGSAEPP	LTAVRPEGSE	P..PGFPTSG	PRRRPGCSRK	NRTSHCRLG	
Consensus		-----	-----	-----	-----	-R-----	
NOC0-R2	451	ACQAPPDSRG	PALSAGLPSP	LLCLLLLVPH	HL~~~~~	~~~~~	491
NOC0-R3		PIRAPSGVQQ	ASSGTALGAP	LLAWILGLAV	TLR~~~~~	~~~~~	
NOC0-R1		QAGSGGGCTG	DSEGSALPS	LTCSLTPLGL	ALVLWTVLGP	C	
Consensus		-----	-----	L-----	-L~~~~~	~~~~~	

3/6
FIG. 2

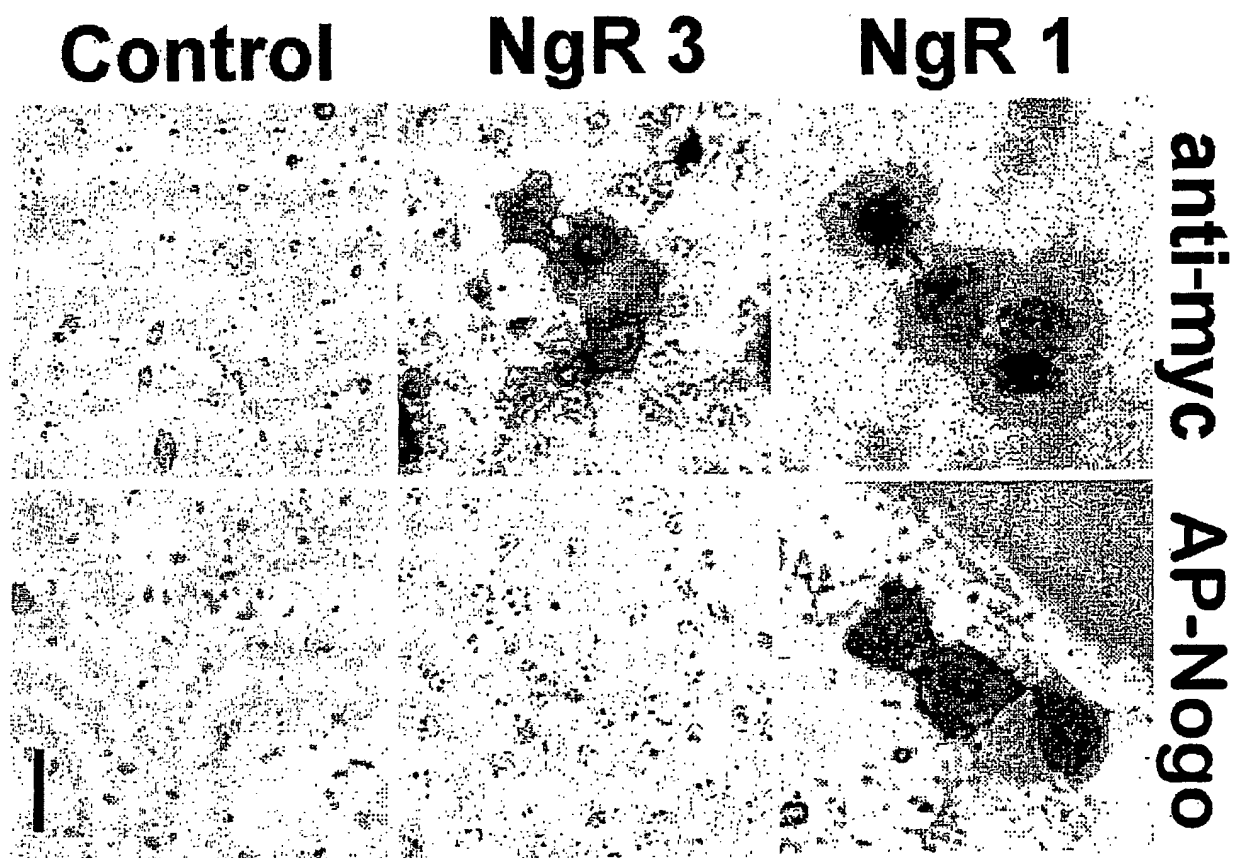


FIG. 3, cont.

	LRR 6	LRR 7	250
Human NOGO-R1	GNRISSVPER	AFRGLHSLDR	LLLHQNRVAH VHPHAFRDLG RLMTLYLFAN
Murine NOGO-R1	GNRIPSVPEH	AFRGLHSLDR	LLLHQNHVAR VHPHAFRDLG RLMTLYLFAN
Murine NOGO-R3	GNKLWSLQGG	IFRGLVNLDR	LLLHENQLQW VHKAFHDLH RLTTFLFN
Human NOGO-R3	GNKLWSLQGG	IFRGLVNLDR	LLLHENQLQW VHKAFHDLR RLTTFLFN
Human NOGO-R2	GNRLRLLTEH	VFRGLGSLDR	LLLHGNRLQG VHRAAFRLG RLTTLYLFAN
Consensus	GN-----	-FRGL--LDR	LLLH-N--- VH--AF--L- RL--L-LF-N
	LRR 8	LRR CT	300
Human NOGO-R1	NLSALPTEAL	APLRALQYLR	LNDNPWVDC RARPLWAWLQ KFRGSSSEVP
Murine NOGO-R1	NLSMLPAEVL	MPLRSIQYLR	LNDNPWVDC RARPLWAWLQ KFRGSSSEVP
Murine NOGO-R3	SLTELQGDCL	APLVALEFLR	LNGNAWDCG RARSLWEWLR RFRGSSSAVP
Human NOGO-R3	SLSELQGECL	APLGALEFLR	LNGNPWDGG RARSLWEWLR RFRGSSSAVP
Human NOGO-R2	SLASLPGEAL	ADLPSELEFLR	LNANPWACDC RARPLWAWFQ RARVSSSDVT
Consensus	-L--L-----L	--L--L--LR	LN-N-W-C-C RAR-LW-W-- --R-SSS-V-
	LRR 8	LRR CT	350
Human NOGO-R1	CSLPQRLAGR	DLKRLAANDL	QGCATATGPY HPIWTGRATD EEPLGLPKCC
Murine NOGO-R1	CNLPQRLADR	DLKRLAASDL	EGCAVASGPF RPIQTSQITD EELLSLPKCC
Murine NOGO-R3	CATPELRQGG	DLKLLRVEDF	RNCTGPVSP. HOIKSHTLTT SDRAARKEHH
Human NOGO-R3	CVSPGLRHGG	DLKLLRAEDF	RNCTGPASP. HOIKSHTLTT TDRAARKEHH
Human NOGO-R2	CATPPERQGR	DLRALREADF	QACP.PAAP. TRPGSRA... .RGNSSSNH
Consensus	C--P-----	DL--L---D- --C-----P-	-----
	LRR 8	LRR CT	400
Human NOGO-R1	QPDAAADKASV	LEPGRPASAG	NALKGRVPPG DSPPGNGSGP RHINDSPFGT
Murine NOGO-R1	QPDAAADKASV	LEPGRPASAG	NALKGRVPPG DTPPGNGSGP RHINDSPFGT
Murine NOGO-R3	PSHGASRDKG	HPHGHPPGSR	SGYK..... .KAGKNC'TSH RNRNQISKVS
Human NOGO-R3	SPHGPTRSKG	HPH.....GPR	PGHR..... .KPGKNC'TNP RNRNQISKAG
Human NOGO-R2	.LYGVA.EAG	AP...PADPS	TLYR..... .DLPA..... .EDSRGR
Consensus	-----	-----	-----

	451	Putative GPI Signals		490
Human NOGO-R1	AGSGGGGTGD	SESGALPSL	TCSLTPLGLA	LVLWTVLGP
Murine NOGO-R1	AGSGASGTGD	AESGALPAL	ACSLAPLGLA	LVLWTVLGP
Murine NOGO-R3	ASSGTALG	APLLAWILGL AVTLR ~~~~
Human NOGO-R3	ASSASSLG	ASLLAWTLGL AVTLR ~~~~
Human NOGO-R2	SRGPALSA	GLPS PLLCL LLLVPHHL ~~~~
Consensus	-----	-----	-----	-----

SEQUENCE LISTING

<110> BIOGEN, INC.
YALE UNIVERSITY
STRITTMATTER, STEPHEN M.
CATE, RICHARD L.
SAH, DINAH W.Y.

<120> NOGO RECEPTOR HOMOLOGS

<130> A116PCT

<140>

<141>

<150> 60/238,361

<151> 2000-10-06

<160> 16

<170> PatentIn Ver. 2.1

<210> 17

<211> 1260

<212> DNA

<213> Homo sapiens

<400> 1

```

atgctgcccc ggctcaggcg cctgctgcaa gctcccgctt cggcctgcct cctgctgatg 60
ctcctggccc tgcccctggc ggccccagc tgcccctgc tctgcacctg ctactcatcc 120
ccgcccaccg tgagctgcca ggccaacaac ttctcctctg tgccgctgtc cctgccaccc 180
agcactcagc gactcttctt gcagaacaac ctcacccgca cgtgcgggoc aggcaccttt 240
gggtccaacc tgctcaccct gtggctcttc tccaacaacc tctccaccat ctacccgggc 300
actttccgcc acttgcaagc cctggaggag ctggacctgc gtgacaaccg gcacctgcgc 360
tcgctggagc ccgacacctt ccagggcctg gagcggtgc agtcgctgca tttgtaccgc 420
tgccagctca gcagcctgcc cggcaacatc ttccgaggcc tggtcagcct gcagtacctc 480
tacctccagg agaacagcct gctccacctc caggatgact tgttcgcgga cctggccaac 540
ctgagccacc tcttccctcca cgggaaccgc ctgcggctgc tcacagagca cgtgtttcgc 600
ggcctgggca gctgggaccg gctgctgctg cacgggaacc ggctgcaggg cgtgcaccgc 660
gcggccttcc gcggcctcag ccgctcacc atcctctacc tgttcaaca cagcctggcc 720
tcgctgcccc gcgaggcgct cgcgcacctg ccctcgctcg agttcctgcg gctcaacgct 780
aacccttggg cgtgcgactg ccgcgcgcgg ccgctctggg cctgggtcca gcgcgcgcgc 840
gtgtccagct ccgacgtgac ctgcgccacc ccccgaggc gccagggccg agacctgcgc 900
gcgctccgcg agggcgactt ccaggcgtgt ccgcccgcgg caccacgcg gccgggcagc 960
cgcgcccgcg gcaacagctc ctccaaccac ctgtacgggg tggccgaggg cggggcgccc 1020
ccagccgatc cctccaccct ctaccgagat ctgcctgccg aagactcgcg ggggcgccag 1080
ggcggggagc cgcctactga ggacgactac tgggggggct acgggggtga ggaccagcga 1140
ggggagcaga tgtgccccgg cgtgcctgc caggcgcgcc cggactcccc aggccctgcg 1200
ctctcggccg ggctccccag cctctgtctt tgccctcctg tcctggtgcc ccaccacctc 1260

```

<210> 2

<211> 420

<212> PRT

<213> Homo sapiens

<400> 2

```

Met Leu Pro Gly Leu Arg Arg Leu Leu Gln Ala Pro Ala Ser Ala Cys
  1                   5                   10                   15

```

```

Leu Leu Leu Met Leu Leu Ala Leu Pro Leu Ala Ala Pro Ser Cys Pro
  20                   25                   30

```

Met Leu Cys Thr Cys Tyr Ser Ser Pro Pro Thr Val Ser Cys Gln Ala
 35 40 45
 Asn Asn Phe Ser Ser Val Pro Leu Ser Leu Pro Pro Ser Thr Gln Arg
 50 55 60
 Leu Phe Leu Gln Asn Asn Leu Ile Arg Thr Leu Arg Pro Gly Thr Phe
 65 70 75 80
 Gly Ser Asn Leu Leu Thr Leu Trp Leu Phe Ser Asn Asn Leu Ser Thr
 85 90 95
 Ile Tyr Pro Gly Thr Phe Arg His Leu Gln Ala Leu Glu Glu Leu Asp
 100 105 110
 Leu Gly Asp Asn Arg His Leu Arg Ser Leu Glu Pro Asp Thr Phe Gln
 115 120 125
 Gly Leu Glu Arg Leu Gln Ser Leu His Leu Tyr Arg Cys Gln Leu Ser
 130 135 140
 Ser Leu Pro Gly Asn Ile Phe Arg Gly Leu Val Ser Leu Gln Tyr Leu
 145 150 155 160
 Tyr Leu Gln Glu Asn Ser Leu Leu His Leu Gln Asp Asp Leu Phe Ala
 165 170 175
 Asp Leu Ala Asn Leu Ser His Leu Phe Leu His Gly Asn Arg Leu Arg
 180 185 190
 Leu Leu Thr Glu His Val Phe Arg Gly Leu Gly Ser Leu Asp Arg Leu
 195 200 205
 Leu Leu His Gly Asn Arg Leu Gln Gly Val His Arg Ala Ala Phe Arg
 210 215 220
 Gly Leu Ser Arg Leu Thr Ile Leu Tyr Leu Phe Asn Asn Ser Leu Ala
 225 230 235 240
 Ser Leu Pro Gly Glu Ala Leu Ala Asp Leu Pro Ser Leu Glu Phe Leu
 245 250 255
 Arg Leu Asn Ala Asn Pro Trp Ala Cys Asp Cys Arg Ala Arg Pro Leu
 260 265 270
 Trp Ala Trp Phe Gln Arg Ala Arg Val Ser Ser Ser Asp Val Thr Cys
 275 280 285
 Ala Thr Pro Pro Glu Arg Gln Gly Arg Asp Leu Arg Ala Leu Arg Glu
 290 295 300
 Ala Asp Phe Gln Ala Cys Pro Pro Ala Ala Pro Thr Arg Pro Gly Ser
 305 310 315 320
 Arg Ala Arg Gly Asn Ser Ser Ser Asn His Leu Tyr Gly Val Ala Glu
 325 330 335
 Ala Gly Ala Pro Pro Ala Asp Pro Ser Thr Leu Tyr Arg Asp Leu Pro
 340 345 350
 Ala Glu Asp Ser Arg Gly Arg Gln Gly Gly Asp Ala Pro Thr Glu Asp
 355 360 365
 Asp Tyr Trp Gly Gly Tyr Gly Gly Glu Asp Gln Arg Gly Glu Gln Met

370 375 380

Cys Pro Gly Ala Ala Cys Gln Ala Pro Pro Asp Ser Arg Gly Pro Ala
 385 390 395 400

Leu Ser Ala Gly Leu Pro Ser Pro Leu Leu Cys Leu Leu Leu Leu Val
 405 410 415

Pro His His Leu
 420

<210> 3
 <211> 1383
 <212> DNA
 <213> Mus sp.

<400> 3

atgtcttggc agtctggaac cacagtgaca caatctccc tgcaggctgc tcaggctctca 60
 ggggtgctgtg tggaattgct gctgttgctg ctgctggag agctacctct ggggtgggtgt 120
 tgtcctcgag actgtgtgtg ctacctgctg cccatgactg tcagctgcca ggcacacaac 180
 tttgctgcca tcccggagg catcccagag gacagtgagc gcatcttcct gcagaacaat 240
 cgcatacctt tctccagca gggccacttc agccccgcca tggtcaccct ctggatctac 300
 tccaacaaca tcactttcat tgcctccaac accttcgagg gctttgtgca tctggaggag 360
 ctagaccttg gagacaaccg acagctgcga acgctggcac ccgagacctt ccaaggcctg 420
 gtgaagcttc acgccctcta cctctataag tgtggactga gcgccctgcc cgcaggcatc 480
 tttggtggcc tgcacagcct gcagtatctc tacttgcagg acaaccatat cgagtacctc 540
 caagatgaca tctttgtgga cctgggtcaat ctcagtcact tgtttctcca tggtaacaag 600
 ctatggagcc tgggccaagg catcttccgg ggcctggtga acctggaccg gttgctgctg 660
 catgagaacc agctacagtg ggttcaccac aaggctttcc atgacctcca caggctaacc 720
 accctcttct tcttcaacaa cagcctcact gagctgcagg gtgactgtct ggccccctg 780
 gtggccttgg agttccttcg cctcaatggg aatgcttggg actgtggctg ccgggcacgt 840
 tccctgtggg aatggctgag aaggttccgt ggctctagct ctgctgtccc ctgctgcacc 900
 cccgagctgc ggcaaggcca ggatctgaag ctgctgaggg tggaggactt ccggaactgc 960
 acaggaccag tgtctctca ccagatcaag tctcacacgc ttaccacctc tgacagggct 1020
 gcccgcaagg agcaccatcc gtcccatggg gcctccaggg acaaaggcca cccacatggc 1080
 catccgcctg gctccaggtc aggttacaag aaggcaggca agaactgcac cagccacagg 1140
 aaccggaacc agatctctaa ggtgagctct gggaaagagc ttaccgaact gcaggactat 1200
 gccccgact atcagcacaa gttcagcttt gacatcatgc ccaccgcag acccaagagg 1260
 aaggccaagt gtgctcgag gacccccatc cgtgccccca gtgggggtgca gcaggcatcc 1320
 tcaggcacgg cccttggggc cccactcctg gcctggatac tggggctggc agtcactctc 1380
 cgc 1383

<210> 4
 <211> 461
 <212> PRT
 <213> Mus sp.

<400> 4

Met Ser Trp Gln Ser Gly Thr Thr Val Thr Gln Ser Pro Val Gln Ala
 1 5 10 15

Ala Gln Val Ser Gly Cys Cys Val Glu Leu Leu Leu Leu Leu Ala
 20 25 30

Gly Glu Leu Pro Leu Gly Gly Gly Cys Pro Arg Asp Cys Val Cys Tyr
 35 40 45

Pro Ala Pro Met Thr Val Ser Cys Gln Ala His Asn Phe Ala Ala Ile
 50 55 60

Pro Glu Gly Ile Pro Glu Asp Ser Glu Arg Ile Phe Leu Gln Asn Asn
 65 70 75 80

4

420 425 430
 Pro Ser Gly Val Gln Gln Ala Ser Ser Gly Thr Ala Leu Gly Ala Pro
 435 440 445
 Leu Leu Ala Trp Ile Leu Gly Leu Ala Val Thr Leu Arg
 450 455 460

 <210> 5
 <211> 473
 <212> PRT
 <213> Homo sapiens

 <400> 5
 Met Lys Arg Ala Ser Ala Gly Gly Ser Arg Leu Leu Ala Trp Val Leu
 1 5 10 15
 Trp Leu Gln Ala Trp Gln Val Ala Ala Pro Cys Pro Gly Ala Cys Val
 20 25 30
 Cys Tyr Asn Glu Pro Lys Val Thr Thr Ser Cys Pro Gln Gln Gly Leu
 35 40 45
 Gln Ala Val Pro Val Gly Ile Pro Ala Ala Ser Gln Arg Ile Phe Leu
 50 55 60
 His Gly Asn Arg Ile Ser His Val Pro Ala Ala Ser Phe Arg Ala Cys
 65 70 75 80
 Arg Asn Leu Thr Ile Leu Trp Leu His Ser Asn Val Leu Ala Arg Ile
 85 90 95
 Asp Ala Ala Ala Phe Thr Gly Leu Ala Leu Leu Glu Gln Leu Asp Leu
 100 105 110
 Ser Asp Asn Ala Gln Leu Arg Ser Val Asp Pro Ala Thr Phe His Gly
 115 120 125
 Leu Gly Arg Leu His Thr Leu His Leu Asp Arg Cys Gly Leu Gln Glu
 130 135 140
 Leu Gly Pro Gly Leu Phe Arg Gly Leu Ala Ala Leu Gln Tyr Leu Tyr
 145 150 155 160
 Leu Gln Asp Asn Ala Leu Gln Ala Leu Pro Asp Asp Thr Phe Arg Asp
 165 170 175
 Leu Gly Asn Leu Thr His Leu Phe Leu His Gly Asn Arg Ile Ser Ser
 180 185 190
 Val Pro Glu Arg Ala Phe Arg Gly Leu His Ser Leu Asp Arg Leu Leu
 195 200 205
 Leu His Gln Asn Arg Val Ala His Val His Pro His Ala Phe Arg Asp
 210 215 220
 Leu Gly Arg Leu Met Thr Leu Tyr Leu Phe Ala Asn Asn Leu Ser Ala
 225 230 235 240
 Leu Pro Thr Glu Ala Leu Ala Pro Leu Arg Ala Leu Gln Tyr Leu Arg
 245 250 255
 Leu Asn Asp Asn Pro Trp Val Cys Asp Cys Arg Ala Arg Pro Leu Trp

260										265					270				
Ala	Trp	Leu	Gln	Lys	Phe	Arg	Gly	Ser	Ser	Ser	Glu	Val	Pro	Cys	Ser				
		275					280					285							
Leu	Pro	Gln	Arg	Leu	Ala	Gly	Arg	Asp	Leu	Lys	Arg	Leu	Ala	Ala	Asn				
		290				295					300								
Asp	Leu	Gln	Gly	Cys	Ala	Val	Ala	Thr	Gly	Pro	Tyr	His	Pro	Ile	Trp				
		305			310					315					320				
Thr	Gly	Arg	Ala	Thr	Asp	Glu	Glu	Pro	Leu	Gly	Leu	Pro	Lys	Cys	Cys				
				325					330					335					
Gln	Pro	Asp	Ala	Ala	Asp	Lys	Ala	Ser	Val	Leu	Glu	Pro	Gly	Arg	Pro				
			340					345					350						
Ala	Ser	Ala	Gly	Asn	Ala	Leu	Lys	Gly	Arg	Val	Pro	Pro	Gly	Asp	Ser				
		355					360					365							
Pro	Pro	Gly	Asn	Gly	Ser	Gly	Pro	Arg	His	Ile	Asn	Asp	Ser	Pro	Phe				
		370				375					380								
Gly	Thr	Leu	Pro	Gly	Ser	Ala	Glu	Pro	Pro	Leu	Thr	Ala	Val	Arg	Pro				
		385			390					395					400				
Glu	Gly	Ser	Glu	Pro	Pro	Gly	Phe	Pro	Thr	Ser	Gly	Pro	Arg	Arg	Arg				
				405					410					415					
Pro	Gly	Cys	Ser	Arg	Lys	Asn	Arg	Thr	Arg	Ser	His	Cys	Arg	Leu	Gly				
			420					425					430						
Gln	Ala	Gly	Ser	Gly	Gly	Gly	Gly	Thr	Gly	Asp	Ser	Glu	Gly	Ser	Gly				
		435				440						445							
Ala	Leu	Pro	Ser	Leu	Thr	Cys	Ser	Leu	Thr	Pro	Leu	Gly	Leu	Ala	Leu				
		450				455					460								
Val	Leu	Trp	Thr	Val	Leu	Gly	Pro	Cys											
		465			470														

<210> 6

<211> 440

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Consensus
sequence

<220>

<221> MOD_RES

<222> (3)..(4)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (6)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (9)..(10)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (12)..(13)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (15)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (18)..(25)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (27)..(29)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (31)..(34)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (36)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (39)..(40)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (42)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (44)..(50)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (52)..(59)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (62)..(63)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (66)..(69)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (71)..(74)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (76)..(80)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (83)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (87)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (90)..(91)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (94)..(96)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (98)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (101)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (104)..(105)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (107)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (109)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (111)..(112)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (114)
<223> Variable amino acid

<220>

<221> MOD_RES
<222> (116)..(117)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (119)..(122)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (124)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (127)..(128)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (136)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (138)..(141)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (143)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (146)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (148)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (151)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (154)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (162)..(170)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (175)..(176)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (184)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (186)..(189)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (192)..(193)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (196)..(197)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (199)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (202)..(203)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (205)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (208)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (210)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (212)..(213)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (215)..(218)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (221)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (223)..(224)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (226)..(227)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (232)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (234)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (236)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (238)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (243)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (246)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (248)..(251)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (253)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (257)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (259)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (261)..(262)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (264)..(267)
<223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (269)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (272)..(273)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (275)..(277)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (279)..(281)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (283)..(287)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (289)..(290)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (292)..(328)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (330)..(341)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (344)..(346)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (348)..(399)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (401)..(428)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (431)..(439)
 <223> Variable amino acid

<400> 6
 Cys Pro Xaa Xaa Cys Xaa Cys Tyr Xaa Xaa Pro Xaa Xaa Thr Xaa Ser
 1 5 10 15

13

355 360 365
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 370 375 380
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg
 385 390 395 400
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 405 410 415
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 420 425 430
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 435 440

<210> 7
 <211> 66
 <212> PRT
 <213> Homo sapiens

<400> 7
 Arg Ile Tyr Lys Gly Val Ile Gln Ala Ile Gln Lys Ser Asp Glu Gly
 1 5 10 15
 His Pro Phe Arg Ala Tyr Leu Glu Ser Glu Val Ala Ile Ser Glu Glu
 20 25 30
 Leu Val Gln Lys Tyr Ser Asn Ser Ala Leu Gly His Val Asn Cys Thr
 35 40 45
 Ile Lys Glu Leu Arg Arg Leu Phe Leu Val Asp Asp Leu Val Asp Ser
 50 55 60
 Leu Lys
 65

<210> 8
 <211> 390
 <212> PRT
 <213> Homo sapeins

<400> 8
 Cys Pro Met Leu Cys Thr Cys Tyr Ser Ser Pro Pro Thr Val Ser Cys
 1 5 10 15
 Gln Ala Asn Asn Phe Ser Ser Val Pro Leu Ser Leu Pro Pro Ser Thr
 20 25 30
 Gln Arg Leu Phe Leu Gln Asn Asn Leu Ile Arg Thr Leu Arg Pro Gly
 35 40 45
 Thr Phe Gly Ser Asn Leu Leu Thr Leu Trp Leu Phe Ser Asn Asn Leu
 50 55 60
 Ser Thr Ile Tyr Pro Gly Thr Phe Arg His Leu Gln Ala Leu Glu Glu
 65 70 75 80
 Leu Asp Leu Gly Asp Asn Arg His Leu Arg Ser Leu Glu Pro Asp Thr
 85 90 95

Phe Gln Gly Leu Glu Arg Leu Gln Ser Leu His Leu Tyr Arg Cys Gln
 100 105 110
 Leu Ser Ser Leu Pro Gly Asn Ile Phe Arg Gly Leu Val Ser Leu Gln
 115 120 125
 Tyr Leu Tyr Leu Gln Glu Asn Ser Leu Leu His Leu Gln Asp Asp Leu
 130 135 140
 Phe Ala Asp Leu Ala Asn Leu Ser His Leu Phe Leu His Gly Asn Arg
 145 150 155 160
 Leu Arg Leu Leu Thr Glu His Val Phe Arg Gly Leu Gly Ser Leu Asp
 165 170 175
 Arg Leu Leu Leu His Gly Asn Arg Leu Gln Gly Val His Arg Ala Ala
 180 185 190
 Phe Arg Gly Leu Ser Arg Leu Thr Ile Leu Tyr Leu Phe Asn Asn Ser
 195 200 205
 Leu Ala Ser Leu Pro Gly Glu Ala Leu Ala Asp Leu Pro Ser Leu Glu
 210 215 220
 Phe Leu Arg Leu Asn Ala Asn Pro Trp Ala Cys Asp Cys Arg Ala Arg
 225 230 235 240
 Pro Leu Trp Ala Trp Phe Gln Arg Ala Arg Val Ser Ser Ser Asp Val
 245 250 255
 Thr Cys Ala Thr Pro Pro Glu Arg Gln Gly Arg Asp Leu Arg Ala Leu
 260 265 270
 Arg Glu Ala Asp Phe Gln Ala Cys Pro Pro Ala Ala Pro Thr Arg Pro
 275 280 285
 Gly Ser Arg Ala Arg Gly Asn Ser Ser Ser Asn His Leu Tyr Gly Val
 290 295 300
 Ala Glu Ala Gly Ala Pro Pro Ala Asp Pro Ser Thr Leu Tyr Arg Asp
 305 310 315 320
 Leu Pro Ala Glu Asp Ser Arg Gly Arg Gln Gly Gly Asp Ala Pro Thr
 325 330 335
 Glu Asp Asp Tyr Trp Gly Gly Tyr Gly Gly Glu Asp Gln Arg Gly Glu
 340 345 350
 Gln Met Cys Pro Gly Ala Ala Cys Gln Ala Pro Pro Asp Ser Arg Gly
 355 360 365
 Pro Ala Leu Ser Ala Gly Leu Pro Ser Pro Leu Leu Cys Leu Leu Leu
 370 375 380
 Leu Val Pro His His Leu
 385 390

<210> 9
 <211> 421
 <212> PRT
 <213> Mus sp.

<400> 9

Cys Pro Arg Asp Cys Val Cys Tyr Pro Ala Pro Met Thr Val Ser Cys
 1 5 10 15
 Gln Ala His Asn Phe Ala Ala Ile Pro Glu Gly Ile Pro Glu Asp Ser
 20 25 30
 Glu Arg Ile Phe Leu Gln Asn Asn Arg Ile Thr Phe Leu Gln Gln Gly
 35 40 45
 His Phe Ser Pro Ala Met Val Thr Leu Trp Ile Tyr Ser Asn Asn Ile
 50 55 60
 Thr Phe Ile Ala Pro Asn Thr Phe Glu Gly Phe Val His Leu Glu Glu
 65 70 75 80
 Leu Asp Leu Gly Asp Asn Arg Gln Leu Arg Thr Leu Ala Pro Glu Thr
 85 90 95
 Phe Gln Gly Leu Val Lys Leu His Ala Leu Tyr Leu Tyr Lys Cys Gly
 100 105 110
 Leu Ser Ala Leu Pro Ala Gly Ile Phe Gly Gly Leu His Ser Leu Gln
 115 120 125
 Tyr Leu Tyr Leu Gln Asp Asn His Ile Glu Tyr Leu Gln Asp Asp Ile
 130 135 140
 Phe Val Asp Leu Val Asn Leu Ser His Leu Phe Leu His Gly Asn Lys
 145 150 155 160
 Leu Trp Ser Leu Gly Gln Gly Ile Phe Arg Gly Leu Val Asn Leu Asp
 165 170 175
 Arg Leu Leu Leu His Glu Asn Gln Leu Gln Trp Val His His Lys Ala
 180 185 190
 Phe His Asp Leu His Arg Leu Thr Thr Leu Phe Leu Phe Asn Asn Ser
 195 200 205
 Leu Thr Glu Leu Gln Gly Asp Cys Leu Ala Pro Leu Val Ala Leu Glu
 210 215 220
 Phe Leu Arg Leu Asn Gly Asn Ala Trp Asp Cys Gly Cys Arg Ala Arg
 225 230 235 240
 Ser Leu Trp Glu Trp Leu Arg Arg Phe Arg Gly Ser Ser Ser Ala Val
 245 250 255
 Pro Cys Ala Thr Pro Glu Leu Arg Gln Gly Gln Asp Leu Lys Leu Leu
 260 265 270
 Arg Val Glu Asp Phe Arg Asn Cys Thr Gly Pro Val Ser Pro His Gln
 275 280 285
 Ile Lys Ser His Thr Leu Thr Thr Ser Asp Arg Ala Ala Arg Lys Glu
 290 295 300
 His His Pro Ser His Gly Ala Ser Arg Asp Lys Gly His Pro His Gly
 305 310 315 320
 His Pro Pro Gly Ser Arg Ser Gly Tyr Lys Lys Ala Gly Lys Asn Cys
 325 330 335
 Thr Ser His Arg Asn Arg Asn Gln Ile Ser Lys Val Ser Ser Gly Lys

340 345 350
 Glu Leu Thr Glu Leu Gln Asp Tyr Ala Pro Asp Tyr Gln His Lys Phe
 355 360 365
 Ser Phe Asp Ile Met Pro Thr Ala Arg Pro Lys Arg Lys Gly Lys Cys
 370 375 380
 Ala Arg Arg Thr Pro Ile Arg Ala Pro Ser Gly Val Gln Gln Ala Ser
 385 390 395 400
 Ser Gly Thr Ala Leu Gly Ala Pro Leu Leu Ala Trp Ile Leu Gly Leu
 405 410 415
 Ala Val Thr Leu Arg
 420

<210> 10
 <211> 17
 <212> PRT
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Consensus
 sequence

<220>
 <221> MOD_RES
 <222> (3)..(4)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (6)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (9)..(10)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (12)..(13)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (15)
 <223> Variable amino acid

<400> 10
 Cys Pro Xaa Xaa Cys Xaa Cys Tyr Xaa Xaa Pro Xaa Xaa Thr Xaa Ser
 1 5 10 15

Cys

<210> 11
 <211> 50
 <212> PRT
 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Consensus
sequence

<220>
<221> MOD_RES
<222> (2)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (4)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (6)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (11)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (14)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (16)..(19)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (21)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (25)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (27)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (29)..(30)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (32)..(35)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (37)
<223> Variable amino acid

```

<220>
<221> MOD_RES
<222> (40)..(41)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (43)..(45)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (47)..(49)
<223> Variable amino acid

<400> 11
Asn Xaa Trp Xaa Cys Xaa Cys Arg Ala Arg Xaa Leu Trp Xaa Trp Xaa
 1          5          10          15

Xaa Xaa Xaa Arg Xaa Ser Ser Ser Xaa Val Xaa Cys Xaa Xaa Pro Xaa
          20          25          30

Xaa Xaa Xaa Gly Xaa Asp Leu Xaa Xaa Leu Xaa Xaa Xaa Asp Xaa Xaa
          35          40          45

Xaa Cys
 50

<210> 12
<211> 196
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Consensus
sequence

<220>
<221> MOD_RES
<222> (2)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (5)..(6)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (8)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (10)..(16)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (18)..(25)
<223> Variable amino acid

<220>

```

<221> MOD_RES
<222> (28)..(29)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (32)..(35)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (37)..(40)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (42)..(46)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (49)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (53)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (56)..(57)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (60)..(62)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (64)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (67)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (70)..(71)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (73)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (75)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (77)..(78)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (80)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (82)..(83)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (85)..(88)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (90)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (93)..(94)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (102)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (104)..(107)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (109)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (112)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (114)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (117)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (120)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (128)..(136)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (141)..(142)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (150)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (152)..(155)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (158)..(159)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (162)..(163)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (165)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (168)..(169)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (171)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (174)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (176)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (178)..(179)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (181)..(184)
<223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (187)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (189)..(190)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (192)..(193)
 <223> Variable amino acid

<400> 12
 Arg Xaa Phe Leu Xaa Xaa Asn Xaa Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15
 Phe Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Trp Xaa Xaa Ser Asn Xaa
 20 25 30
 Xaa Xaa Xaa Ile Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa Leu Glu
 35 40 45
 Xaa Leu Asp Leu Xaa Asp Asn Xaa Xaa Leu Arg Xaa Xaa Xaa Pro Xaa
 50 55 60
 Thr Phe Xaa Gly Leu Xaa Xaa Leu Xaa Leu Xaa Leu Xaa Xaa Cys Xaa
 65 70 75 80
 Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Phe Xaa Gly Leu Xaa Xaa Leu Gln
 85 90 95
 Tyr Leu Tyr Leu Gln Xaa Asn Xaa Xaa Xaa Xaa Leu Xaa Asp Asp Xaa
 100 105 110
 Phe Xaa Asp Leu Xaa Asn Leu Xaa His Leu Phe Leu His Gly Asn Xaa
 115 120 125
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Arg Gly Leu Xaa Xaa Leu Asp
 130 135 140
 Arg Leu Leu Leu His Xaa Asn Xaa Xaa Xaa Xaa Val His Xaa Xaa Ala
 145 150 155 160
 Phe Xaa Xaa Leu Xaa Arg Leu Xaa Xaa Leu Xaa Leu Phe Xaa Asn Xaa
 165 170 175
 Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Leu Ala Xaa Leu Xaa Xaa Leu Xaa
 180 185 190
 Xaa Leu Arg Leu
 195

<210> 13
 <211> 1176
 <212> DNA
 <213> Homo sapiens

<400> 13
 gagggcatcc ccgtggacag cgagcgcgctc ttcctgcaga acaaccgcat cggcctcctc 60
 cagcccggcc acttcagccc cgccatggctc accctgtgga tctactcgaa caacatcacc 120

```

tacatccacc ccagcacctt cgagggttc gtgcacctgg aggagctgga cctcggcgac 180
aaccggcagc tgcggacgct ggcacccgag accttccagg gcctggtgaa gcttcacgcc 240
ctctacctct acaagtgtgg gctcagcgcc ttgccggccg gcgtctttgg cggcctgcac 300
agcctgcagt acctctacct gcaggacaac cacatcgagt acctccagga cgacatcttc 360
gtggacctgg tcaacctcag ccacctgttt ctccacggca acaagctgtg gagtctgggc 420
ccgggcacct tccggggcct ggtgaacctg gaccgtcttt tgctgcacga gaaccagctg 480
cagtgggtcc accacaaggc attccacgac ctccgcaggc tgaccaccct ctccctcttc 540
aacaacagcc tctcggagct gcagggtgag tgcctggccc cgctgggggc cctggagttc 600
ctccgcctca acggcaaccc ctgggactgt ggttgtcgcg cgcgctccct gtgggaatgg 660
ctgcagaggt tccggggctc cagctccgct gtcccctgtg tgtcccctgg gctgcggcac 720
ggccaggacc tgaagctgct gagggccgag gacttccgga actgcacggg accagcgtcc 780
ccgcaccaga tcaagtcaca cagctcacc accaccgaca gggccgcccg caaggaacac 840
cactcacccc acggccccac caggagcaag ggccaccgcg acggcccccg gcccgccac 900
aggaagccgg ggaagaactg caccaacccc aggaaccgca atcagatctc taaggcgggc 960
gccgggaaac agggccccga gctgccagac tatgccccag actaccagca caagttcagt 1020
tttgacatca tgcctacggc ccggcccaag aggaagggca agtgtgcccg caggaccccc 1080
atccgtgccc ccagcgggt gcagcaggcc tctcggcca gttccctggg ggcctccctc 1140
ctggcctgga cactggggct ggcggtcact ctccgc 1176

```

<210> 14

<211> 392

<212> PRT

<213> Homo sapiens

<400> 14

```

Glu Gly Ile Pro Val Asp Ser Glu Arg Val Phe Leu Gln Asn Asn Arg
  1             5             10             15

```

```

Ile Gly Leu Leu Gln Pro Gly His Phe Ser Pro Ala Met Val Thr Leu
      20             25             30

```

```

Trp Ile Tyr Ser Asn Asn Ile Thr Tyr Ile His Pro Ser Thr Phe Glu
      35             40             45

```

```

Gly Phe Val His Leu Glu Glu Leu Asp Leu Gly Asp Asn Arg Gln Leu
      50             55             60

```

```

Arg Thr Leu Ala Pro Glu Thr Phe Gln Gly Leu Val Lys Leu His Ala
      65             70             75             80

```

```

Leu Tyr Leu Tyr Lys Cys Gly Leu Ser Ala Leu Pro Ala Gly Val Phe
      85             90             95

```

```

Gly Gly Leu His Ser Leu Gln Tyr Leu Tyr Leu Gln Asp Asn His Ile
      100            105            110

```

```

Glu Tyr Leu Gln Asp Asp Ile Phe Val Asp Leu Val Asn Leu Ser His
      115            120            125

```

```

Leu Phe Leu His Gly Asn Lys Leu Trp Ser Leu Gly Pro Gly Thr Phe
      130            135            140

```

```

Arg Gly Leu Val Asn Leu Asp Arg Leu Leu Leu His Glu Asn Gln Leu
      145            150            155            160

```

```

Gln Trp Val His His Lys Ala Phe His Asp Leu Arg Arg Leu Thr Thr
      165            170            175

```

```

Leu Phe Leu Phe Asn Asn Ser Leu Ser Glu Leu Gln Gly Glu Cys Leu
      180            185            190

```

```

Ala Pro Leu Gly Ala Leu Glu Phe Leu Arg Leu Asn Gly Asn Pro Trp
      195            200            205

```

Asp Cys Gly Cys Arg Ala Arg Ser Leu Trp Glu Trp Leu Gln Arg Phe
 210 215 220
 Arg Gly Ser Ser Ser Ala Val Pro Cys Val Ser Pro Gly Leu Arg His
 225 230 235 240
 Gly Gln Asp Leu Lys Leu Leu Arg Ala Glu Asp Phe Arg Asn Cys Thr
 245 250 255
 Gly Pro Ala Ser Pro His Gln Ile Lys Ser His Thr Leu Thr Thr Thr
 260 265 270
 Asp Arg Ala Ala Arg Lys Glu His His Ser Pro His Gly Pro Thr Arg
 275 280 285
 Ser Lys Gly His Pro His Gly Pro Arg Pro Gly His Arg Lys Pro Gly
 290 295 300
 Lys Asn Cys Thr Asn Pro Arg Asn Arg Asn Gln Ile Ser Lys Ala Gly
 305 310 315 320
 Ala Gly Lys Gln Ala Pro Glu Leu Pro Asp Tyr Ala Pro Asp Tyr Gln
 325 330 335
 His Lys Phe Ser Phe Asp Ile Met Pro Thr Ala Arg Pro Lys Arg Lys
 340 345 350
 Gly Lys Cys Ala Arg Arg Thr Pro Ile Arg Ala Pro Ser Gly Val Gln
 355 360 365
 Gln Ala Ser Ser Ala Ser Ser Leu Gly Ala Ser Leu Leu Ala Trp Thr
 370 375 380
 Leu Gly Leu Ala Val Thr Leu Arg
 385 390

<210> 15

<211> 143899

<212> DNA

<213> Homo sapiens

<220>

<221> modified_base

<222> (2044)..(2144)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (6609)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (6625)..(6724)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (14153)..(14252)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (19512)..(19611)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (22595)..(22694)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (27825)..(27924)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (34953)..(35052)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (40783)..(40882)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (49000)..(49099)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (62884)..(62983)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (75528)..(75627)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (87944)..(88043)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (111030)..(111129)
 <223> a, t, c, g, other or unknown

<400> 15
 aagcacatac aggtgacatt acagaactga cagttatgcc aggcactgta cttagcccct 60
 ataccatcct caaacagctg tatgatgtag attgggtatt aaccccatta ataacaaaag 120
 tacagggaac aaagtgactt tccaaaggctc atgccattca aaggagggtg aatcttaggt 180
 tggacgcagg ctgtctgact ctggagtctg aggtgttaat gctgcctcct ccatgggaac 240
 agcccaagtg aaaaacagct gatccactct tcatttactt ggcactctgtg ctaagctggt 300
 ccctgagcca agctctgagc aacagaaaca gaagctctgc attaggagct tgtgagcatg 360
 tcaatgccgg gtaaaggagt gctggaacc gctgggatgg ccgccgagca ctaggccgtt 420
 gaaggtgggc tctgtgtgac tggttcctct acactctggc ctggctgcct gcaggaagaa 480
 gatcaagctg agtgggctgg ccctggacca caaggtgaca ggtgacctct tctacacca 540
 tgtgaccacc atgggccaga ggctcagcca gaaggccccc agcctggagg acggttcgga 600
 tgccttcattg tcaccccagg atgttcgggg cacctcagaa aaccttcctg agagtgagt 660
 tctggtcaag gtgccggcct tgggggatag tgatgggtgg tcctcatatt cagtgcacac 720
 tcatggttga gtattttatt gcacccctct tcagtcctta caacacccca tgatgtaggt 780
 ggggcatgct cctcatttac agatgggcac atcaaagctc agctaacgct ggggaagtta 840
 gattcagggt taccctgctg gattcctggg attggggagg gaggagcttc caaaatgggg 900

acaaggtctc	tgggcctgtc	gggtagctgg	tttccctcagg	gccccttgca	acctctgagc	960
ttattgcatc	aggtgcagcc	aggcccgtga	gcctcctggc	aggggtcctc	cacacctggc	1020
tgtcttttgc	cccctgctgg	tcacaggagg	agctgcagca	cctgcctggg	ctgcttctca	1080
ggaggggtaca	tgaagatccc	aggaccgcca	gtcccatgat	aagtgggaagg	agctccttgg	1140
agtcaggagc	gggagttgag	gagtttgagt	cctgctctcc	agttataggc	tatgtgacct	1200
gtgtagatca	cctaaccctg	ctcttgattt	ccttacctct	taaactagca	ctaaaagcac	1260
cccacaaact	gtaaagttag	ttgtgatgat	tgaatgacac	catgggtgtg	gaagctcttt	1320
gtaaagtgc	aaacgggtgtg	cagtttgagg	gtggttacc	ccagtgccga	ttctcagagg	1380
gcaacatggc	taagggcacg	agctggagtt	aggctgacct	gctgcttcga	gccctgtgag	1440
cttgagcaag	tcattttaact	tcctgagctg	cagtttcctc	atcagtaaaa	tgtgataagg	1500
ataggggtgt	tgtaagattt	tattaaatgg	ggtaataaat	gtcaagtatg	tagcccatag	1560
tgagtgcctc	agagtttttt	tcttttgttt	ctttccccc	cgccccgaga	tggagcctta	1620
ctctgttgcc	caggctggag	tgcagtggca	tgatcttgge	tcactgcaac	ctccgcctcc	1680
cgggttcaag	caattctcct	gcctcagcct	cccaaatagc	tgggactaca	ggcgtgcacc	1740
accatgctcg	gctaattttt	gtatctttag	tagagacggg	gtttcaccat	gttggccagg	1800
ctggctctcga	actcctgacc	tcagtatgct	cctgcctcag	ccccgaaaag	ttttgggatt	1860
acaagtgtga	gcccccggtg	cctgccaggt	tttttttttt	tttttttttt	tgtaaaaacac	1920
ccacagggtga	ttgctgttgc	ctgggctgga	gtgcggtagt	gcaatcatag	ttcactgcag	1980
ccttgacctc	ctgggctcaa	gtgatcctcc	tgccctcagc	tcctgagtag	ctgggaatac	2040
aggnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	2100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	2160
agacagggtt	ttcccgaatgt	tggccaaggc	tggtctaaaa	ctcccaacct	caggtgatcc	2220
accacctca	gcctcccaaa	gtactgggat	tcaggcgctg	agacaccgtg	cccagccagg	2280
aggcttattt	tcttgataaa	ttaccagtc	tcagggtattt	ctctacagcg	atgcaagaac	2340
agcctaatac	atccaggctc	agcatcagtg	gacccagggtg	ggagagctta	agatgtcaag	2400
gtctgaatgc	cgcttccaca	cacctttggg	acctagggac	tcctctctct	tttctttttt	2460
cagtagaaga	tgttatcttc	tcctttctct	gaccagtagt	tgggtgatgt	ttcagagata	2520
gtttttcagt	caagatatat	ttcagtggct	tcactgagcc	caagttccct	cgctctcta	2580
ggactttatt	tccttgtttc	tagaagagg	ataacacata	ttttctaagg	tggttgtgag	2640
attaaggagg	ctgggtaccg	gtgggtgcata	aggacaggat	agagcaatgg	tgagaccact	2700
caaaaagcga	aaagttgacc	tgcgagggtg	acacttatca	aatcagcaca	cagtgggagt	2760
ggaaggaatg	tcctcatca	gttacaatat	ttggagagtg	caagttatag	aaaaccagc	2820
cctggccggg	cgcggtgggt	catgcctata	atcccagcac	tttgggaggc	tgaggcagg	2880
ggatcacgag	gtcaggagtt	caagaccagc	ctgaccaacg	tgggtgaaacc	ccacctctac	2940
taaaaataca	aaattagctg	ggcgtgggtg	tgtgtgcctg	taatctgagc	tactcaggag	3000
gctgaggcac	gagaatcact	tgaaaccggg	aggtggagtt	tgcaagtgagc	cgagatcgca	3060
ccactgcact	ccagcctggg	caacagagcg	agactccatc	tcaaaccgaa	aaaaaaaaag	3120
aaagaaaacc	cagctctaac	tggcttaaac	agtaagaaga	tctattatat	tatccatctc	3180
aggcagcagc	aagcccagag	gtaggggact	ccaaggttgg	ttgatccagg	gcttaacgat	3240
gtcatcaaag	acccaggttc	ttctgtctc	ggcacctctg	tctgcagggc	cagcttcac	3300
ctaagccaga	ttgttcttgt	cttgattaca	agttggctgc	tgggccagca	gacgctgcct	3360
gcctccctgt	tcactcttcag	aagtagaaag	tggcccttcc	ccagtcatgg	aatgaaagag	3420
tttccctttc	gtctgggatt	gcttaggtcc	acccacctga	agccaatgac	tgtaaccagg	3480
aaggtaatat	acactgattg	tcttaagtca	gggttctctga	gccagtcctg	ggcaaggagt	3540
gtgatactgt	catgattgtc	ttgggctcat	cagggcagct	ctgcagatga	gatcaaaact	3600
caagctacat	tattctgaac	agtgggaagt	aggaaagaga	catttttgga	gatacaaaac	3660
acaatgtcta	tcccatatcc	ctaggtccag	gtcacagtgt	cttggttggg	catcaaattg	3720
agaaaaagaa	agactgtcca	tccatttatc	tacctattca	tctggttttt	gatttttttt	3780
aaatttttatt	ttaagacatt	ctcactctgt	caccagact	ggagtgcagt	ggtttgatca	3840
tggctcatgg	cagcctcaac	ctcccaggct	caagtgaacc	tcccatgctc	aagtgatcct	3900
cctacctcag	cctoccaaagt	agctagaact	aaaggtgcac	gccaccacgc	tcagtttaatt	3960
tttgcatttt	ttgtagagat	ggggtttcgt	catgatgccc	atgctagtct	ggaattcctg	4020
aactcaagca	atatgcctgc	ctttgcctcc	caaaatgctg	ggattgtagg	catgagccac	4080
tgtcctggc	tcactctgtt	aataatttat	gaaacaacta	ctgggtgctg	agcacggggc	4140
caggggctgg	agatctagca	gggaccaggc	agatctctgc	caagtcggtg	gtttctttaa	4200
ggttttgctc	ataattcccc	ttttcttttc	tctttcgttt	ttttcttttt	ctttctttct	4260
ttctttcttt	tttttttttt	gagacagagt	ctcactctgt	taccaggtct	ggagtgcagt	4320
ggtgcgatct	cagctcactg	caacctctgc	ctcctgggtt	caagcgattc	tcctgcctca	4380
gcctcccag	tagctgggac	tacaggcgcc	tgccaccatg	cccggttaat	ttttgtgttt	4440
ttagtagaga	ctgggtttca	ccatattggc	caggctggtc	ttgaactcct	gacctgtgta	4500
tcggcccgct	tcggcctccc	acagtgcctg	gattacaggc	gtgagccacg	gcgcccagcc	4560
agtttccctt	ttcaatgagg	cctccctgac	ctccatactc	tactcctcca	cctggcccac	4620
tcagctctac	tttttcttcc	ccatagcact	caagacctcc	taacatacta	cgtaagttat	4680
ttatttacta	ggcttactgt	gtattgtctg	tcttccctca	ctagaatgta	aactccatga	4740

gaatagaaat	ttttgccttt	ttatttagtg	tggtgtctgc	agccctggc	ttagtccttg	4800
gcatacaaca	gtcactccac	ccacagttgc	tgaataagtg	actaaaggct	cctgccctca	4860
tattgttatg	agggagtggt	catgttggtta	gagaaaaatc	tgaggcaca	taaaatttta	4920
tagagtttaa	gttttctttt	ttaagcaatc	cacgaattgg	ggtagtttca	gaggtagttt	4980
ttcagtcag	acgtatttca	atggcttcac	tgagcccaag	ttctttcacc	tctctaggac	5040
tttatttcct	tatttctaga	acggggataa	cacatagttc	ataaggcagt	tatgagagta	5100
agggagctgg	tatgggggtga	tgcataagga	caggatagag	cagtgggtgag	accgctcaga	5160
tgacaaagcg	tcagagacca	gtattttacga	cggaaatgtg	gaagcatgat	aaagaaatta	5220
tttgggctgg	gcacaatgac	tcacaactaa	taaaactttg	ggaggccaag	gtgggaggat	5280
cacttgactt	gcagaaggtc	aaggctgcag	tgagctgtga	ttttgccact	gcactccagc	5340
ctgggtcaaca	gagtgagacc	ctggctcgaa	acgttatttg	attgggttaca	gttatacagt	5400
tgcccttattt	gggtctattcc	atttgaaagt	tcctagttct	ataattttta	gtttgttggt	5460
tggttctgat	tggttaagct	taagttttgt	tttcttttaa	tacagttaag	tgccccataa	5520
tgacattttg	gtcaaggaca	gaccacatat	acagtgggtg	ttccataaga	ttataatgga	5580
gctgaaacat	tcctattgtc	tatggcgtag	tggtcctgat	gttgtagcgc	aatgcattag	5640
ttatatgttt	gtggcaatgc	tggtgtaaac	acactactg	cactgccagt	gatataaaag	5700
aatagcacat	acagttatat	atagtagata	atatctgata	atgataatac	ataactatat	5760
tactggttta	tatatattact	atattattta	tctttatttt	atttttgaga	cagagtctca	5820
ttctgtcacc	caggctggag	tgagtggtgc	cgatcttggt	tcaccgcaac	ctccgcttcc	5880
tggtttcaag	tgattctcct	gcctcagttc	cctgagtagc	tggtgattaca	ggtgtgcacc	5940
atgacaccct	gctaatatgt	tttgattttt	tagtagagat	gggtgtttcac	catgttggtc	6000
aggctggctg	tgaactactg	acctcaagtg	atcacccgc	ctcggcttcc	caaagtgtct	6060
ggattacagg	cgtgagccac	cacgcatggc	ctattttata	ttatttttaga	gtgtacgcct	6120
tatacttata	aaaaaaagct	aactgtcaaa	cagcctcggg	caggctccttc	aacagatatt	6180
ccagaagaca	ttgttatcat	aggagatgac	agctccgtgc	atattattgt	ccctgaaaac	6240
cttctagttg	ggaagtggaa	gacagtgata	ttgatgatag	gacccagtgt	aggcctaggc	6300
taattgtgtg	gtttgtgtct	ttgcttttaa	caagaaagtt	taaaaagtta	aaataaaata	6360
caaaaattttt	taaatagaaa	aaagctgccc	aggaacaatg	gctcacacct	gtaatccac	6420
cattcgggga	ggccaagggtg	ggtggattgc	ttgagctcag	gagttcaaga	ccagcctggg	6480
caacatgggtg	aaaccccatc	tctacaaaaa	atacaaaaat	tagccgggtg	tggtggcatg	6540
cggctatagt	tccagctaata	cgaggggctg	aggtgggagg	atcactgggg	gggaggtggt	6600
tgaggctgna	gtgagctgtg	attgnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	6660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	6720
nnnnatattc	ttaaaaaaat	ttttttttat	ttttgagaca	gaattttctc	cttgttgccc	6780
aggctggagt	gcaatggcgc	tatctcagct	cagggaaccc	tccacctcct	gggttcaage	6840
gattctcctg	ccttagcctc	ccaggtacag	gcgcccgcga	ccatgctcgg	ctaatttttg	6900
tatttttagt	agagatgggg	tttcacatg	ttgtccaggc	tggtcctgaa	atcctgcctc	6960
aggatgatcca	ccccctcggg	cctcccaaag	tgctggaatt	tacaggcgtg	agccactgtg	7020
cctggctccc	tttacatttt	tttaaattta	atgttaattt	tttaattttt	aattttctcat	7080
atatatatat	ttttaagact	agccaagtga	agcagtggga	gtggaaaagg	aactggtttt	7140
gatcaatagg	tgtaaacacc	actgcactgg	gaccagccta	ttttacattc	ctgttagcag	7200
tgatgagggg	tcactttctt	tgtagcctca	acaatatgtg	tcgttgccca	tctttttttt	7260
tttttttttt	tttttttttg	agatggagtc	tcaactgtgt	gcctaggctg	gaatgcaatg	7320
gcatgatctc	agctcactgc	aacctccgcc	tccagggttc	aagtgtattc	tgtgtctcag	7380
cctcctaggt	agatgggatt	acaggcgtcc	accaccacgc	ccggctaatt	ttttgtattt	7440
tcagtagaga	tggtgtttca	ccatgttggt	caggttggtt	tcgaactcct	gacctcaagt	7500
gatccgcccc	cctcggcctc	ccaaaagtgt	gggtattacag	gcatgagcca	ccgcgcccgg	7560
cctgcccatc	ttttttttgt	tatagccatc	ctagtggatg	taaagttttt	ttgtgatttt	7620
gatttgtgtt	tcctactga	tcaatgatgt	tgagcatctt	ttcctgtgct	tattggcttt	7680
tggtatatct	ttggagaaag	gtctattcag	gtcctttgcc	cactttaaaa	ttagggttatc	7740
tttctattac	tgagatgtaa	gagttcttta	tgttctagat	ataagtctcc	tacatatgat	7800
ttgtaaaaat	tttccttcca	ttattgggtt	gtcctttcact	ttccttttgg	gtccttttagt	7860
gcacaacagt	ttttaaatatt	gaagtccaat	tttctatttt	tctcttttgc	cacttgatct	7920
ttgggtgtcat	gtttaaggaa	ctattgccta	atctcagggtc	acaaagattt	acacctgtgt	7980
ttccttcttt	ccttccttcc	ttccttccct	ccttccttcc	ctccctccct	ctctccctcc	8040
ctccctctct	ccctccctcc	ctccttccct	ctcctccctc	ctccctccct	ccttccctcc	8100
ttccttccct	ccttccctcc	ttccttccct	ccttcccttg	tccttctgac	ggaatcttgc	8160
tctgtcacc	aggctggagt	gtagtggcac	gatcttgggt	cactgcaacc	tctgcctcct	8220
gggttcaagc	aattctcctg	cctcagcctc	ctgagtagct	gggactacag	gcacacacca	8280
ccatgcccag	ctaatttttg	tatttttagt	agagacgggg	tttcaccaca	ttggccagga	8340
tggtttogag	ctcctgacct	cgtgatccac	cgcccttggc	ctcccaaagt	gctgggattg	8400
caggtgtgag	ccaccatgcc	cggcctgtgt	tttcttagag	ttttgtagtt	ttagctctta	8460
tagttagatc	cttgatccat	tttgagtga	ttttgtatat	agtgtagat	atccacctgg	8520
tgttgtaaat	tgcccagaag	tgggtatgct	tctaaatctg	gctgttaggg	attactagag	8580

gtgaccaaag	tgaatttttt	ctttgtttct	tttttttttt	ggagacagag	tctccgtcac	8640
ccaggctgga	gtgcaatggc	ttcatcttgg	ctcagtgcaa	cctctgcctt	ctggtttcaa	8700
gcagttctcc	tgccctcagac	tccctgagtag	ctgggtattac	aggcgtgtac	cacccatgctt	8760
ggctaatttt	tgtattttta	gtaaagatgc	agtttcacct	gttggccagg	cttttctgga	8820
actcccggcc	tcaagtgatc	catctgcctc	tacctcccaa	agtgtctgga	ttacaggtgg	8880
gagccaccgt	gcccagtcct	tttctcagaa	tttatttggt	tttttttggt	ttgtttcatt	8940
tttgagatag	ggctctcactc	tgctcagctag	gcaggagtcc	agtgggtgtga	tcattgctgc	9000
agccttgaac	ttctggactc	acgtgatctt	cccacctcag	cctcctgagt	agctaggatt	9060
acaggcatgt	gcttccacac	ctggctaatt	ttttaatttt	ctaggactta	tttgtccatt	9120
cttgcaaaagc	agggtacaac	atgcctatct	ctacctacct	ctcttccctt	caagggactc	9180
cagccaaaat	ccttgaggct	ctcgggctga	ctgtgggtgc	tggtgcctga	tctgcctcag	9240
tcatgctgca	tgatcaaaaag	tgtccgtttt	ctgcttcttg	gaactttatt	cactttgggt	9300
gtcagctcttc	ctctgcagtg	tcccaagaac	acagaattag	accaggaatc	tgtgttgcca	9360
tagtgtgtgg	aaagaggcag	acttccaact	ccgctatgtg	ctgttgggtg	attgaagctt	9420
aattttcttt	ctatctttct	ttcttttctt	ttcttttttt	ttttttggag	atggaatctc	9480
gctctgttgc	ccaggctgga	gtgcagtggt	gcgatctcac	ctcactgcaa	cctccgcctc	9540
ccaggttcaa	gcgattctcc	tgccctcagcc	ctcgtgagtag	ctgggattac	aggtgcatgc	9600
caccatgccc	gcctaatttg	tgtaatttta	gtagaacacag	tgtttcacca	tatttggtcag	9660
gctggtctcg	acctcctcac	ctcaggtgat	ccaccgcctt	tggcctccca	aagtgtcggg	9720
attacaggcg	tgagccaccg	tgccctggcac	ttaattttct	taatacctca	attaccccat	9780
atggtaaaat	gggactagta	atccatacct	tatagcgctg	ttgtgaaaat	gaaatgaggg	9840
taagcagata	aaatttcaga	ctacggatgg	gattgttact	acattctgaa	cctggctttg	9900
ctgttatttg	ctatgtgacc	ttatcttctc	tggtactcca	ttctttccaa	gtctataaaa	9960
caaagtggac	aattgtcaac	ctttcttcca	aagagcaatg	atttaaggat	caaagtatgt	10020
catttaacaa	aaatatgaag	agctcaacaa	atgaggaact	cattattatt	attacaatta	10080
ttattatttt	agaaataggg	tcttgttctc	ttgcctagge	tggagtccag	tggtataaac	10140
acagctcaat	gcattctcag	cctcctggat	acaagtgate	ctcatgtctc	atccccctaa	10200
gtagctggga	ccacaggcat	gtaccaccac	gcacggctaa	ttttttattt	tttattttta	10260
ttttttgaga	cagtcttgct	ttgtcgcaca	gactggagtg	cagcagcgca	atcaccgctc	10320
actgcaacct	ccgcctcctg	ggttcaagtg	attctgctgc	ctcaacctcc	caagtagctg	10380
ggattacagg	cctgtgccac	catgcccgcc	taattttttt	gtattttttg	taaagacggg	10440
gtttcaccat	gttgcccagg	ctgatctaga	acccttgccc	tcaagtgate	cccctttctt	10500
ggcctcctaa	agtgtcagga	ttacaggcgt	gagcctctgc	acctggcctc	ggctaatttt	10560
ttattttttg	tagagacagg	ttctcactat	gttgccaggg	ctggtcttga	actcctgggc	10620
tcaagtgate	ttcccacctc	agcctcccaa	agtgtcgaga	ttacagatgt	gagccactgt	10680
gcctggcctg	gaactcatta	ttgaagcatt	cactagtatc	aactttgggg	ttacctggcc	10740
acatcctctg	acctacctat	aagggtatca	cagctaaccg	agcctctggt	tctcagaatt	10800
taggcagaag	cagttcaatt	tatcacaaac	tactctatat	ccagcataag	tgcccaaata	10860
aaacaattgc	taaagttctt	taggcattta	ctgtttgtta	gttagatatt	tagtcctcac	10920
tacaaatctg	tgatacagg	attattttta	tttaacccat	tttatagaag	agaaacctga	10980
agctcagaga	tgctaagtaa	cttgtgcaag	gtcacacagc	tagtaaataa	agggcagagt	11040
aaagatttag	tttcacattg	gactccagaa	cctttctact	gggactcatg	ggaatagtgt	11100
ggatgtccct	gaccttcagt	ggcccagggc	tctcctgggg	gaatccagcc	atagacaaga	11160
caccagcgag	agcccaatcc	taagattttg	ttgttttggt	tttgagacaa	ggctcactc	11220
tgtcaccaga	ctggagtgc	caatgctcac	tgcaaccttg	actacaggtg	cacaccacca	11280
ctcaagcaat	cctcccacct	cagcctcctg	agtagcttgg	ctattatttt	ttgagacagg	11340
cacctgacta	attttaaaat	tttattttaat	taattactta	ctcagctcat	tgcgctcctc	11400
gtatcacttt	gtcacccaag	ctggactgca	atgggtgtgg	agttgcaggg	actgcagggt	11460
acctcccagg	ttcaagtgat	cctcccacct	cagcctctgg	gatggggtct	cactatgttg	11520
tgcgccacta	tgctcagcta	atgtttttat	tttttgtata	tgtcttggcc	tcccaaagtg	11580
ccagggtctag	tctcaaaactc	ttggactcaa	gcgatcctcc	gtgtttttgc	tgaatgtgac	11640
ccgggattac	aggcataaac	caccacaccc	aacccttaag	aaggaacact	gagcagggag	11700
catgtcagag	gcaggaaagg	gaagcatcat	gggggttagga	tgtatgtgtg	acaattctca	11760
acaaagaaaa	tgggatcatt	ttgtgagtg	tcgctgtgtg	cccagatgag	ataatggagc	11820
gagccagcct	ctcaggtggt	tgagaccaca	gtccccattt	tgaacccggc	tcttccagac	11880
ctcagagagt	ttctgcagca	cagctagtgg	aattagaatt	gtttaccttc	atgattttac	11940
tccaggtgct	tcacaacccat	cccaaaccct	gtcatttgca	tccttttgaa	tcagtctgct	12000
catttccctt	tgccatagct	agtgttattt	acttaataat	aatatcgggt	ttgctgttac	12060
taaaaaaaaa	tagcttcatt	ctaaagtgt	atattcttgg	acatatatat	gtacgtatat	12120
ccacccccac	acgttataca	tatacatgta	tgtttctaatt	tttttttaga	tggagctctc	12180
acgtgtatcg	ttttttgtta	tgaggtgtg	tgatttcggc	tcactggaac	ctctgcctcc	12240
ctctgtagcc	caggctggag	gcctcagcct	ctggagtagc	tgggattaca	ggcaccacc	12300
tgggttcaag	cgattctcct	gtatttttag	tagagacagg	gtttcacccat	gttggccagg	12360
actacacccg	gctaattgtt					12420

tggtgcttga	actcctgatc	tcaagtgatc	cacctgcttt	ggcttcccaa	agtgtctggga	12480
ttataggtgc	gagctactgc	ggctggccaa	tgtatgtttt	taatacacat	tcaaataacg	12540
aataactatg	aaacctgaaa	aactgctcca	tgttacttcc	tgaacctatc	ttgagtgtc	12600
acatgtctgt	cataccacat	attgggaaac	actgctttcc	ctggcttcca	agccagctt	12660
aatcactgtc	ccatcctatg	cttcgcttta	tttgtctata	aatgttgggg	ttgggggttg	12720
atgccaaaga	ccttttctgt	tgctattaac	atggacacag	ctctaagagg	tcttggcatc	12780
ttgggctggc	tctcctttta	gttcagaatt	tggattttta	tccaactact	cagagtgtatc	12840
aagccttcct	tatgaatgaa	ctcgttggtc	aaactcataa	aaggctgtatc	gataaaacag	12900
gaatgaatgt	atgaattgac	actaagtcat	tagcatttca	cgggaatgga	ttctccgtta	12960
gtggaagagc	acatgtcctt	tctggcactg	atgtgtgctt	gggaaactta	ctgagctaac	13020
tggcccatgt	aacacagagg	ccctttgggtg	cagtggaaaa	ctgttgactt	tggagattat	13080
ccttagtctg	aatctgagcc	tgctgttaag	aagctggcta	actgaattgc	tttgcttctt	13140
ggacccttac	cattttataaa	atggggacca	ttgtactcac	ccttttaggt	tattgcatgg	13200
attaaatggg	attctctata	gaaaatattg	gcacaaagta	ggtgtaaaatt	tgacgcgtag	13260
tgggattgtt	tgtgagggaa	attgtcattt	gattatcaaa	gacttaggag	caggaacagt	13320
gtctaattca	gggactgcaa	atggaaatgc	cagctgaggc	caggcatttg	ctaataattg	13380
ggtaaaagcag	ggcaggtgta	gaatagcaat	gtctgggaat	taaaagagag	gtgaggacgt	13440
gtatgacctt	gagaaggcaa	gccctggcaa	aaggggatgg	cctccactca	gctacagtca	13500
tgccctagatc	ttctaacttt	ttatttttat	ttttattttt	tgagacggag	tcttgctctg	13560
tcacccaggc	tggagtgcag	tggcgcgatc	tccgctcact	gcaagctccg	cctcccgggt	13620
tcacgccatt	ctcctgcctc	agcctcccaa	gtagctggga	ctacggggcg	ccaccaccat	13680
gcccggctaa	tttttttttt	gtatttttag	tagagatggg	gtttcaccgt	gttagccagg	13740
atggtctcga	tctcctgact	ttgtgattta	ccctccttgg	cctcccaaag	tgctgggatt	13800
acaggcttga	gccaccgcac	ctggccgatc	ttctaacttt	ttaaagagaa	gcaagacatc	13860
tggattttta	tgtgataact	cctgattttta	aactggcacc	caattataat	ttacaacact	13920
ataagggtca	acattgccag	cagagcaaaa	catgggtggg	ggcaactgct	ggtcacccgt	13980
gtgcagcctc	tgggtctaaaa	tcactcttgt	atttcttctt	gctttacgca	ttgtcccagc	14040
acagtgcctg	tgtatagtaa	atatccagta	agtgggtgta	gaatgaataa	accaatgcag	14100
ataaacctgt	agagaggccg	ggcacagtgt	ctcatgtctg	taatctcagc	acnnnnnnnn	14160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14280
gggtagatca	cctgaggtca	ggagttcaag	accagcctgg	ccaatatagt	gaaaccccgt	14340
ctctacaaaa	ataaaaaaat	tatctgggca	tgtattgcag	tgccctaat	cccagctact	14400
cgggaggctg	aggccggaga	attgcttgaa	cctgggaggc	ggaggttgta	gtgagccgag	14460
atcatgccat	tgactccag	cctaggtgac	ggagcaagat	tctgtctcaa	aaaaaaaaaa	14520
aaaaaaaaag	aaaaaagaaa	agaaaaagaa	acaatgaatg	agtgtgaggc	tcatggtagt	14580
attggttcct	gagagttagc	aaccttattg	gtcatcccag	ccacgaagtg	aaatggtagc	14640
cctggcttgg	gccaatgaat	gaggaagaat	aatggcaaat	gggggtctat	gcctccacc	14700
ctcaccacta	gggaggtctc	aagcttgaaa	tccagtgacc	agggttttag	gtcctggacc	14760
tggccagtcc	tcctacagtc	aagtagataa	gtggagggtt	tgggtccgtg	ggctacggag	14820
atagtgtatca	aggccgttac	tctgcaatca	gactcagaaa	tggcctctca	gttacttctc	14880
cattttgtggg	tcttttgtaa	gagcagagaa	gaggaaggaa	tttaggtcct	ctcaccctct	14940
gggctgcctg	tcctgtctcc	ctgagccatg	gagggctggg	gtggaatatg	gggaataaat	15000
ctgtactttt	tttttttttt	tttttttgaga	cagagtctcg	ctccgtcgcc	caggctggag	15060
tgccgtggcg	tgatctctgc	tcacagcagc	atctgcctcc	cgggttcaag	ttattcttcc	15120
acctcagtct	cctgagtagc	tgggattaca	ggtgcccacc	accacgcccg	gctaattttt	15180
gtattttttag	tagagacagg	gtttcactgt	gttgggcagg	ctgggtctcaa	atacctgacc	15240
tcagggtgatc	caccgcgaca	tgccctccaa	agtgtctgaa	ttacaggcat	gagccaccgt	15300
gcccggctcct	accaatctgc	acatttttaat	tgacaagggt	caccctccac	tcatgtgcc	15360
ggcatagtct	tgagaagcat	cccacaagg	tgccctctgag	ttcaccctga	caagtccact	15420
agctcttggc	agagacatct	ggcaaattca	aggcttgaga	catgctggcc	tctcttttaa	15480
gtgcagcaaaa	ttttgtctag	agcttgggtca	gttaaaattt	tgatgttttg	ttttgcatta	15540
atttcaattt	ttaagaaatg	ttgcattaaa	atgttattta	tcttgaatag	ttaaatttctt	15600
agtgtccctt	taatttctta	gtgtgtctga	gttgagagcc	tcccctgcct	gattctagtc	15660
cagaccctgg	ggtgacagaa	gactggtggg	agatgggagg	tgaggagggg	agtgttgggt	15720
ggagagcatg	atctacagag	tgctggagag	actctgtatg	gagcttttca	tgctgcctgt	15780
ttgccagccc	tgaagctatg	ccttgagggt	gggcaagggt	gcatatccta	gatcagagat	15840
cctcaactgg	ggccattttt	ctccccagag	gacatttggg	aacatgtgga	gacatttttg	15900
atcatctcgc	ggggtgggga	gaggggctac	tgacatctgg	tgagttagaga	ccagagggac	15960
cattaaactt	tctacaacgc	ccaggacagc	ccctccacaa	taaagagtta	tttgacctca	16020
catattaata	gcacaaagt	gaggaacctt	gatctagatc	cacagcacag	aagaaaggat	16080
gtagattttt	cacacattaa	agatgagaaa	gcttgtgcct	gtaatccctg	tgactcagga	16140
ggctgtggca	ggaggattac	ttgagcccag	gaattcaggg	ttacagtga	ctatcatcgc	16200
agcactgcac	tccagcctgg	gtgacagagc	aagattttgt	ctcttaaaaa	aaaaaaagat	16260

gaggacaggc	acagtggctc	atgcctgtaa	tcccagcatt	ttgggaggcc	gaagtgggtg	16320
gatcacgagg	tcaggagttc	aagaccagcc	tggccagcat	agtgaacccc	catctctact	16380
aaaaatacaa	aaaattagcc	agctacttgg	gaggctgagg	caggagaagc	gcttgaaccc	16440
gggaggtgga	gcttgcagtg	agccaaaatc	ttgccattgc	actccagcct	gggcgacaga	16500
gcaagactcc	gtctcaaaaa	gaaaaaaaaa	aaagatgaga	aagaggaaag	gagagaaaaa	16560
agagagagag	gaaagaaaga	gagaagggtt	tggagtcaaa	aagacttaga	aattccagtt	16620
cttccacttc	ccatggaaac	ttggcaagtt	gccttctctc	tttctctgaa	tctcacattt	16680
tgctctgtg	aagtaggggt	ggtacctggt	ggagatgatg	cggagatgag	ggtgaggggt	16740
gtgttgacac	ctatgcccct	aggatgggtg	agagcttggg	agcactgaac	ctccctttcc	16800
cctcttgttt	cttcccccca	ttgtctccca	ccagctccct	gggatctcca	cttcactctc	16860
tgggattcca	ccagcaggag	gctactcctg	gagttaaggc	gtgttggtca	gactggggca	16920
ttttaggggg	cataaataat	aattatgcct	ggacaatgga	cataacatct	agggccttct	16980
gaagcaaac	aggggtgtgg	gtacccaaac	aaggcagtag	gccccaggag	gcaggtccct	17040
gcagtcccag	cagagagcag	ggcacagggt	tgagaagact	gagcaaacct	cattatcagc	17100
tcctttgtcc	cccactctgt	cctggagcaa	tcattctggc	ctcttcccac	ttccccaaaa	17160
accaggtata	aaggctgctt	ctggccctctg	aagccagagg	cactgagagt	ggaggtctca	17220
gactcttgga	aggtgagttc	ttttctggct	gcccaggcag	gaccagtgtg	ggccctggga	17280
agaagcagca	cctcataggg	caaacacgta	ggaggcctgt	ccttaggaac	atcatagcta	17340
agcagacctg	tccccgcagg	ggcaggagtc	tgggctaagg	gtgatactgg	agagcagcaa	17400
cggagactgg	aagacaaatg	aaatttggtg	cctgagttat	ccctcccacc	attccttttc	17460
tagactctcc	agctcagggt	ctgttcatgg	caagaggaga	aagcaatctt	gtttgctctt	17520
taatcaaaaa	atataaaaaa	tattccctct	atactatgtg	ccaggggcta	tactagacac	17580
acaaagacag	ccccaaagag	gacggtggag	tagtgtcctc	gctaaaaagac	agtagatatg	17640
caatgcctct	tgctcctctc	ctttctcctg	ctgggaacag	tttctgctct	tcatctgggt	17700
aagtctctcc	cttccctcct	catgcgtctt	tccctttttt	cctttttcct	acactcccct	17760
cccccgctt	ttatttgac	tcatgaggcc	aggaccacag	ccttccctct	ttagctgata	17820
cagctcatct	ccggtaaagt	atcacttggg	ctcagaactg	taacctggaa	ctttctcttt	17880
tttgtttgat	ttttttttgt	tggtgtgtgt	tttgtttttt	ttttgttttg	ttttttgttt	17940
tgttttgaga	cggagtctcg	ctctgttgcc	caggctggag	tgagtggtcg	cgatctcggc	18000
tcaccacaaa	ctccgcctcc	cgggttcaag	caattcttct	gcctcagcct	cctgagtagc	18060
tgggactaca	ggcacatgcc	accacgcctg	gctaattctt	gtatttttag	tagagatggg	18120
gtttcaccat	atgttgccagg	ctggtctcaa	actcctaacc	ttgtgattcg	cccgcctcgg	18180
cctcccaaag	tgctgggatt	acaggcgtga	gccaccgcac	ccggcaaac	gtaacctgaa	18240
ctttcagaag	gaaaaaccac	ccacctgtta	agatgaagg	ctggtgactg	ccccaggctt	18300
ctcacacgtg	ctttctccca	ccttcaaac	acacactcgt	ggtgtcggcc	agaagtcagg	18360
ttcttgtcca	tttgtgggtg	tgacccgaga	gatctctcct	tacctaacac	caaggaaatc	18420
ctccagctct	gtcttcaggt	ggaattccta	ggaaagctcg	agcgacgttg	ctggagctgt	18480
ccacggtgct	ggaactagga	agctcttgac	ctgatggcag	gttacctctt	cttcccagag	18540
aatgatgccc	cccatctgga	gagcctagag	acacaggcag	acctaggcca	ggatctggat	18600
agttcaaagg	agcaggagag	agacttggct	ctgacggagg	aggtgattca	ggcagagggg	18660
gaggagggtc	aggcttctgc	ctgtcaagac	aactttgagg	atgaggaagc	catggagtcg	18720
gacccagctg	ccttagacaa	ggacttccag	tgccccagg	aagaagacat	tggtgaagtg	18780
cagggaagtc	caaggtgcaa	gatctgccc	tacctattgg	tgcggaactc	taaaactttt	18840
gcagaagctc	aggtaagtag	tagggaggct	actgcggagg	acctggggga	aaagagagta	18900
cattcagctc	tctgttccct	attcatttag	gctagtgtgt	ctcaaagcct	cgcatgcata	18960
agaatcacct	ggagttgttg	ttaaaacaca	gctttctggg	cctcacctgc	acgacttctg	19020
athtagagg	gctgaggtga	agcctgagaa	tttgcattta	caacaaatcc	ccaggtgatg	19080
atgatattgt	tggtctgggg	agaaccaccg	atttaaacaa	aaggcttttg	tgtagaaac	19140
gcctgtgtta	aattctggtt	ctgcctttta	ttagctgtgt	tacctgggca	agttgctttg	19200
cctttcaaa	cttttagcacc	ttcatttgta	aaacgaagat	atatagcacc	aacttcttag	19260
agttgtgggt	agcattaaat	gagataatac	atgaaaaagt	tttggaaatg	tactgggct	19320
gtaataaaat	ctcaataaag	ggtggttata	attattatga	gtattatcat	ttctgttagg	19380
attgtcctga	cagctaatta	agaagcaaaa	gataggatta	agggaggcaa	gtaggtttat	19440
ttttaacctg	aaaagggatg	ccgggctctt	gcctggagac	tcagaaactt	gaaataaatg	19500
agagggaatt	cnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	19560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	19620
gtagacacat	agccagaaca	ctagaagggt	gtggtaggag	tggggattag	aggttccagc	19680
tgagggcaat	ggcacttgca	aaggctttgt	tgaagtggcg	taagtgtgga	ggtggagcat	19740
tcaggaaagg	agagcttcag	cttcagtggt	gctggagtgc	tggtgtgtaa	gagaggtgaa	19800
gatgaggctt	ggaggctggg	cagattttgc	tccaaaagag	cttgggtgaac	tgtagtaagg	19860
agtttggtat	ttctcctact	aaggacaaca	gcaaacattt	gaagagttta	aatcggtcag	19920
tgacaatgac	acgtttgcgt	tttgggtgct	cactcgagct	gccagccagg	tagacagtgg	19980
cagaagatgg	aagataaagc	actaaagggt	gatgaggcag	gaagccagtg	aggagagaaa	20040
ggggacgatg	tgagtgcacg	taaatacatt	gttgggtgct	tattgtgtgc	taagctctgt	20100

gctaaattct	tcacgtgtat	tatttcagct	aatccatcta	acaactctgt	aaggcaggta	20160
caatcgttcc	cagctgaaga	agctgaggct	ctcaaaagct	agtaacttgc	ctaagttcat	20220
gcagcatgca	agttgtccag	ccaggattct	aacttagaca	ccagaggcca	cttttaacca	20280
ctgctctagg	actgggggaa	atgggtcccta	gtgagatag	tgctcagttt	catatttcat	20340
tcaacaatat	tggttgccctg	ctacatgtga	agagctgtgg	aaagcgccca	aagtgaagta	20400
gataccctatg	agcaagtggg	atgggggtgg	agtggacagt	aggagggtg	gaacacacat	20460
aaaagggtat	aagaaataac	aattaggccg	gccagggtg	gtggctcacg	cttttaatcc	20520
cagcactttg	ggaggccgag	gagggtggat	cacttgaggc	caggagtttg	agaccagccc	20580
ggccaacatg	gtaaaacccc	atctctacta	aaaatgcaaa	aattagctgg	gctggtgggtg	20640
cacgcctgta	atcccagcta	cttgggaggc	tgaggcacga	gaatcacttg	aaccaggag	20700
gcagagggtta	cagtgaactg	agattgcacc	actctactcc	agcctgggag	acagagtgtg	20760
accctgtctc	aaaaaaagaa	aacaaaacaa	gtaggtaactt	tctgcatag	ggaggattca	20820
taaactgcta	gtcctcaggt	gcatttttgc	ttatcagttt	taaaaatcag	agaatgtctc	20880
aaagaattag	gatgtcagct	tcttttgaaa	atttgggcca	gaagcgggtg	ctcacgcctg	20940
taatcccagc	actttgggag	gctgaggtgg	gtagatcacc	cgaggtcagg	agttggagac	21000
cagcctgacc	aacatggcga	aacccccgat	ctactaaaaa	tacaaaaatt	agctgggctg	21060
gtgggtgcatg	ccttttagttc	cagctactca	ggatgctgag	gcatgagaat	cacttgaacc	21120
cgggaggcag	gggttacagt	gaaatgagat	tgccaccactg	cactctagcc	tgggagacag	21180
agcaagaccc	tgccctgaaa	aaaagaaaaa	gaaaattttgg	aagatctgac	aacagttgac	21240
ctgcattcct	gctcggcaac	agcctgatgg	tggatgggca	gaggctcagt	tgtctgccaa	21300
acctcccac	actgatgtct	tccctcgctg	tcatcatctg	cttgacatgt	aggcatttgg	21360
tgtgtgcctt	ctgctctggg	tgcccagatg	aattggatgc	tatatgagaa	aacattctgt	21420
aaatgtcttg	tggtaggcaa	cctcaaagat	cactggggcc	tccaatgatc	cctccttctt	21480
ggatattcatg	cctgtgtata	atcctctccc	ttgagtgtgt	actacacctg	gatacttgct	21540
tctaataaac	agaacacagc	aagggtaatg	ggatgctact	tctaagggtta	aattacaaga	21600
gtgtaaagtc	tgtcttgttt	gtttccctct	cttgatcttc	ctctcattct	ctctctctcc	21660
ctctctctca	ctttcttact	gtcttgtcct	tccctttgtt	tactctgatg	aagcaagcta	21720
gcaagcatcc	atggtgtgag	ctgacctatg	aagaggccca	tgtgggtgta	aggaactgag	21780
ggcagcctct	acccagcaag	gaactgagtc	actcatcata	tgggtgagct	tggagacaaa	21840
tccctcccca	cttgagcttt	cagatgacgg	cagccctggc	tgatgctttg	caggcttgtg	21900
agagaccctg	agacagaaca	ctcagctaag	ctatacccta	tctcctgaga	tagagtataa	21960
tacatgtagt	tttaagctac	tatgttttgg	gataatttgt	tactcagcaa	tagataacca	22020
atacatatac	catgtacata	actgtttcag	ttgtctgaga	ctatatattag	tcattttaca	22080
cctacatcaa	gaatgtgtca	ggcaccattc	caggtacttg	gaatacatca	attaacagaa	22140
taggtaaaga	ggccaggcat	agggctcaca	tctataatcc	cagcactttg	ggaggccag	22200
gtgggaggac	tgcttgagcc	caggagttag	gaccagcctg	ggtaaaatag	tgagacactg	22260
tctcaactaa	aaaaaaaaaa	aattagttgg	gcacagtggc	acatgcctgt	ggtgccagct	22320
gctcaggagg	ctgaggtggg	aagatcgctt	gagcccagga	gtttgaagct	ccagtgaacc	22380
acggtcacaa	aactgcactc	tagcctgagc	aacagaaaaa	gaccctgtct	caattaaaaa	22440
aaaaaaaaaa	aaaaggaaag	aaagaaaaaa	ataggtaaag	atccttgatt	cttgccctct	22500
tggaacttct	attctagagg	gggatgggtt	ttcacagtag	aagtctgtgt	tgacagcgct	22560
gtttaaagct	ccttcagcat	ctggggaaaa	ggttnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22620
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22680
nnnnnnnnnn	nnnnattttt	tagagatagg	gtcttgctat	gttgcccacc	aggctggtct	22740
tgaaactcctg	ggctcaagca	atcctcctgc	ctcagcctcc	tgagttagctg	ggaatacagg	22800
tgtgcaccac	catgcctggc	ttatttcata	tatatatatt	tttatatata	tgtatatatta	22860
tatatataaa	tatatatata	atttctgtat	ataaataaat	aaatatatat	atatatatat	22920
tttttagagat	agggctcttg	tatgttgacc	accaggtcct	gaactcctgg	gctcaagtga	22980
tccctcctacc	tctgcctttc	aaagtgttgg	gattacaggc	gtgagccatg	gcacctaaact	23040
gagttatttt	taccacacga	agcataggac	atacatccaa	aatgtttctg	agctgagcaa	23100
gagcctggag	gcaagtgaat	ctgaactttc	cctgttttga	agaaaccagt	ctctctccaa	23160
agtcacatag	ttagtgtcac	tccccccaag	aactgcatga	gctgggacaa	tcagagggca	23220
gtggaaggtc	tggggctcag	gggcgcccc	tgctgtctcc	ccagggtctg	tccccttacg	23280
caagagcctc	tgctccccc	ctttcctgtg	gagcctcctc	accatgggca	tgaccagct	23340
gcggatcatc	ttctacatgg	ctgctgtgaa	caagatgctg	gagtaccttg	tgactggtgg	23400
ccaggagcat	ggtgaggcac	cgctgaggcc	cctgggggtt	gggggcacag	gcgggtcacc	23460
ctggctgagc	tccctcacc	atacgtttcc	ctacccacag	agacaaatga	acagcaacaa	23520
aagggtggcag	agacaggtag	ggctatgaaa	gcagggccct	ggctcacgcc	cacccactg	23580
caaccgcctt	ctcagggggc	gggactcctc	taggcctggg	cccacccagg	taaccctttt	23640
gtgggatgta	agagtctggg	ttcagaggaa	ggctattttg	gtgctctctg	gcctccgctg	23700
gaaggggtga	tagtgtccac	tgagtgccag	ttcctgaccc	cactgcccct	cccactctgc	23760
ccagttgggt	tctactcctc	cgtcttcggg	gctctgcagc	tgttgtgcct	tctcacctgc	23820
ccctcattg	gctacatcat	ggactggcgg	atcaaggact	gcgtggacgc	cccaactcag	23880
ggcactgtcc	tcggagatgc	caggtgacct	gcctgtacag	ggatgggtgac	agcaagtgg	23940

caggcagtgc	ttttcatttt	ctctgtgcgt	ttacatccag	cagcttggtg	ctttctccca	24000
agaaccctag	gagatcaggg	gtacctcccc	attttacaga	tgaggaaact	gaggctagga	24060
agggacctgg	cttgcttaat	aataagaata	gctaatgcag	agtgcgtact	gtgcacttgg	24120
caccttgcc	tgtttagtcc	tacaacacct	ctttgaggta	gatgcgttaa	tatcttcatt	24180
ttgcagttga	ggaaaccgag	gtacaggggt	gcacagttag	gtcattcacc	caagatcaca	24240
cagctttcag	tggcagcctc	cagaacctgt	gttataaggg	tacacgctaa	agtcttggtta	24300
gggctagaat	aggtagagtt	ggtatattag	atattttattg	ctgtataaca	aatcacccca	24360
aggcttgcca	ttttaaaaca	acaaacactt	ctcatctcat	acagtttctg	acagtcagaa	24420
atcaggagga	gactcagccg	gctgattctg	agtcacagtc	tctcatgaag	acatagtcag	24480
gctgtcagcc	agggtgcag	tcatctgaag	ggctgactgg	ggttgagaa	tctatgtcag	24540
ttcaattacc	cccattggcct	ctccataggg	ctgctcagga	cacagcacct	gctttccctt	24600
gagcaagaag	gctaagcgac	agagaccccg	tatcttctct	cacataatct	cagacgtagc	24660
ataccatcac	ttctgttacg	ttctattata	ggcacagagc	aacctgata	tactgtggaa	24720
ggagactgga	caaagcaggg	gaataccagg	aggcaggatc	cttgagggct	gtcttggttg	24780
ctggagacca	ccattgaggg	tttttttttt	tttttttatt	gagacagtct	tgctctgtcg	24840
cccaggctgg	agtgcagtg	cacgatctca	gctcactgca	acctctgcct	cccaggttca	24900
agcgattctc	ctgcctcagc	ctcccagta	ctggggttc	acctggagt	cttgaaccca	24960
gattctgtga	ctgcttttgc	tctttttgtg	ttcatccaaa	cagtccctgt	ttatcctaag	25020
aggatgggag	aaagagactg	ggagagaagg	aaatccagtg	gcctccctcc	ctgctagcag	25080
agcctggccc	tggcactgag	ccttcctcct	ctaccctctg	ctcctaattg	tgaggggtccc	25140
ctagcagggc	ccttctgtcc	aggacacatg	ggcgcctgt	cctcacccca	gcctactgac	25200
ctctctcctg	ggctggcctc	agtgccttg	attgtgccgg	agagagggaag	cgctggacag	25260
tcaggccaag	ctgctgtccc	caggagggca	tctgcttatg	tctagggcag	ggacaccttc	25320
ctgaggactt	ctgatgagag	acgggtgtgag	agcttcccac	ttcccacctt	ccttcccctc	25380
cttggttctc	aaaccttcaa	gtgtgcatga	gaatcactta	gtgggggata	tttgtccaaa	25440
tgagatttg	cagatatccc	cgctgagatt	ctgagggccg	agatgaggcc	tgtgaatctg	25500
catgtaaga	aagcaccgcg	tttgatgcgt	gtgtcattgg	gtaggggagc	aacactttga	25560
gaaacattga	gctagagaa	gtgggtttct	atgggtttcc	catagaaaaca	tggatttctg	25620
tgttttctgc	tgcctgaca	tccaaggcac	atctgaagg	ggagggggcca	ggccaagaac	25680
caggaggtcc	tgggaacgta	gaggcagcag	ccagtgaact	cccgtactcc	tcagggacgg	25740
ggttgctacc	aaatccatca	gaccacgcta	ctgcaagatc	caaaagctca	ccaatgccat	25800
cagtgccttc	accctgacca	acctgctgct	tgtgggtttt	ggcatcacct	gtctcatcaa	25860
caacttacac	ctccaggtac	ccaccttcac	tctcccctc	tccctgcctc	ccgaggtccc	25920
tccaaaggga	tggtccatcc	agcacctgcc	ttccaggaag	cgcagttctg	gtcttctgat	25980
ctggatctat	tttccgggtt	ctccaggaag	tgtttctagt	agattgggtt	ggcgaggggg	26040
tgggaattga	ggcccagttg	gcctcttcgc	cctacccctc	cttccctcag	cctccacaca	26100
ctctcctaac	ctcttcactc	tctctttttg	gttttagttt	gtgacctttg	tcttcacac	26160
cattgttcga	ggtttcttcc	actcagcctg	tgggagtctc	tatgctgcag	tgtgagtctg	26220
ttgggctgaa	atgccttcc	gagctttgca	accgtgatca	gagaacccca	gggaagggtt	26280
gggagggcc	caggcatccc	ctaattgcac	tctctctgag	acctctgat	ggcaggagac	26340
tcacttcctt	aaaggcagcc	tatctgtctg	taattgactc	cccctgttgg	agtcttccct	26400
tagaggaagc	tgaaatacct	ggcttgatga	cactttggtt	ctatgtctgc	tgtttgaaac	26460
ggcccccaga	atggcctccc	ctccatgccc	acctgaaga	aatttcccaa	gggcagccat	26520
ttgccttata	attttcctct	tcatgttgga	cagtcaccc	ttgcctctct	ctcctgggtt	26580
cccctgctgg	gcgtgctga	gggactctcc	cctgtgtatg	tgatggagta	acaggacatt	26640
acaataatga	tgacaaaatg	acaaccatta	tcaagtgtct	cgttggtgca	ggcagcaggg	26700
aggatccttg	accatcactc	cctgagttca	gcctcactgc	agcgggtctg	gcagagggca	26760
gctctctttc	cttcatctgc	tcaagccaga	acctggagt	ttccttgatg	tttctctccc	26820
tcacactcca	tgttcactcc	atcctcagta	cagccagcag	cagcttctac	acaccccaaa	26880
tctgacctt	cttgtcacct	ccactgctgc	ctctccagtc	ctagccacca	acatctctag	26940
cctggattat	tgtggcagcc	tttagtctcc	cacatctgcc	ctggccccgc	tgtctcagtc	27000
tatttttaac	acaggggctg	cagtcacctg	tcaggacata	agtctcttca	catcactctg	27060
tggtgtcctg	tctcatctgt	ctcagagtaa	aagccaaagg	ctttactatg	gcctaaaaag	27120
ccctgcaagc	tctggcccca	gcacttcact	cccctctagc	tccccctcct	ccattgttca	27180
ctctgccaca	gccacagtgc	ttcctagtgc	tcgggaagtc	tcaagtgtgt	tccctgcttg	27240
gcatcttttg	atgtactagt	ccctgtttct	agaacattct	tctccagata	tctgcaaggt	27300
gccaatctt	accttctctc	cttcttcagg	tctttccctg	actgtcctct	tctcagtgag	27360
gcctcccttg	gctgtcccat	gtacaattgc	aacctcccta	ctgcccgtt	ctctgcttgg	27420
tttttctcag	cgtttatcac	taacactctg	cctatctctt	gcttattgtc	tgaccgccac	27480
ctgctccatg	ggaatgccac	ctcctcgatg	gcaggaatct	gttgacttgc	ttgatcgtgg	27540
tatctccagc	acatagagca	gtgcctggca	catagtaggt	tctcagctaa	atggttgggtg	27600
acagaataca	gtggacagtc	ctgcgaggtc	aatgccatcc	ctgttattag	tggaggaagt	27660
ggggctcagg	gagtttgagc	cacttgccaa	tatcacacat	acaggaggtg	tgagaaccca	27720
gctcagtgcc	cctgaagttg	gagcatttgc	cctcaaggct	ggggacccaa	gagcccatgc	27780

aaagagcccc	aacgcttaag	caccaccctg	cctggccagc	ggggnnnnnn	nnnnnnnnnn	27840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27900
nnnnnnnnnn	nnnnnnnnnn	nnnncccc	gcgccctggc	cattactttt	aatggcaaaa	27960
accacaatta	cttttgcacc	cacataaata	gttaccatgg	gctgagcatg	gtggctcagg	28020
cctgcaatcc	cagcactttg	ggaggctgag	ccaggcggat	cacttgaggc	caggagtcca	28080
agaccagcct	ggccaacatg	gtgaaacccc	gtctccacta	aaaaatacaa	aaattagctg	28140
ggtgtgtgtg	cgcggtgcctg	taatcccagc	tattcaggag	gcagagggtg	cagttcactg	28200
aaatcatgcc	actgcactcc	agcctgggcg	acagaatgag	actctgtctc	aaaaataaat	28260
aaataaataa	ataaatatatt	accatgtttt	gaccacctgt	tatgtgccaa	ctgtattact	28320
taaaaacacc	catgggaggc	tgggcacagt	ggctcacgcc	tgtaatcgga	cactttggaa	28380
gggcaagcgg	ggaggatccc	ttaaggccag	gagttcaaaa	ccagcctagg	taacacagta	28440
agccctgtct	ctacaaaaaa	taaaaaaatt	aactgggcat	ggtggtgtgt	gcctgtaacc	28500
ccagctcctc	gggaggcaga	gggagagggt	cgcttgagcc	cagcagtttt	aggttgagct	28560
gagccaggac	caagacacta	cactccagcc	tgagtgcag	agcaagacac	tgccctctaaa	28620
caaacaaaca	aacaaaagcg	acctgtgggt	aggtaggaac	aggctcatag	tacagatgag	28680
aaagcagagc	ttggagggct	caagcgatgt	gccaagcaga	ggtccaagcc	gaggtctctc	28740
tgaattccaa	gttaattccg	tctatcatat	caccacagcc	ctctctgccc	cagggagagt	28800
ctctgcccac	tccagccact	cacgtgtaat	tgacttcctc	aggggcagga	aaggcttcga	28860
tgggccaggt	gaggtgagc	ttcagaaaga	taaggcagcc	caggccagac	caggtgaaca	28920
tgatgaccac	gaaggccaca	ccggcatcgt	agatcagctg	tgagaggagg	gggcaggccc	28980
gtgggggaga	ctgcctggcc	ccagacccca	ccaaggtaga	tcccaggcct	cagaggcctt	29040
aaagaagttc	tcttctcccc	ttgtccttgt	gccaattttg	cagatgagga	aaccaagacc	29100
agaagttttg	agtcagactc	agaagaccca	tcatctcttt	ttctttttca	cttgaggccc	29160
cctagagagc	tatgaaatag	tctccacaaa	gcctgaagtt	gctggccact	ggctcaaaat	29220
atctctgaaa	tttccattat	cttaaaaaaa	tacatacatt	tttgcctatg	actccacaaa	29280
cattcatggt	catgttcgca	caaaaaatgtc	catttcatag	tacgtacaaa	ggaaacttag	29340
tgtcttaggt	ttaccggggc	taatcgtggt	tatcctgccc	cttccctggc	cattccccag	29400
gggaaaaggc	aaacccagac	tgctcatgct	cagccttttc	tcacctttcc	caggtcctcc	29460
cacgtgcaac	aactgggggg	gttggggaga	gggaggtgca	agtgtctctc	ccaagggtcc	29520
tcaaccccag	ggcaggtaag	ttctcaattg	aatgagattc	tgtgcaaatg	tgtcagccct	29580
tcttatggaa	gaagctgatg	caccatctgt	cctcttgtcc	tccccatacc	atctgaccag	29640
gataattaat	gtctgtctct	ccctcaggct	cctgtctcaa	cctttttctc	tgcatctttg	29700
gaccttggtg	ccttttctct	cctaggggca	ggacagagct	tcaaaggggc	acacccccaa	29760
atgtgtggag	gtaagatctg	gctcttcaaa	cactacttca	gttgaaaaga	aggggagaact	29820
gcccaccctc	catgcctgcc	caccagaaca	actgatggcc	cccccaacca	tgcgctctct	29880
caaactcctt	tggagacact	gagcaaaagt	accttcttta	gtactctttg	taaagtgcaa	29940
aacggtatgc	agtttggtac	tgcccaccgt	ggagggtgag	gagcatggca	tggctcaaag	30000
ggtcccttga	tatttgacag	aggaaattga	ggcccccatc	ttgactgag	ctaaaacttt	30060
ggtcccttga	cttcgaggta	caccaggttg	acctgtccag	gatccagcct	ggcataaact	30120
cactttgtga	ccttggacca	aaccacccat	cctctctgga	agggttgga	aaatgtggcc	30180
ccaaaggctg	aataaagcca	gagagtcagg	gaccttgaac	gcattgtga	gggctggact	30240
tgattctgta	ggtgaagcta	aaccactgaa	ggtttttcag	cagtgtgtga	gccagttccc	30300
catctgagat	ctttctggaa	gtcacgtgag	tgacagagta	cagagaaaaa	gaatcagagg	30360
cagggagacc	agctgagaaa	gcttgcctgt	gccaggagga	gagggggaag	gcctgcattg	30420
ggatgatgac	agagaaagga	gagcggagaa	gtcagagccc	tgggtcagca	ctagctgctg	30480
ctcactcggc	cccacccggt	tcttgtgtca	agacaaaaag	aaaacccagg	tggcctcata	30540
ccttgattcc	tgggaacgta	atggcagaag	aggcgtaaga	gccaatcatg	agggccatta	30600
acgtggagcg	caggttccca	aacatgtttg	gcagctgagg	agggaagca	gcacccatga	30660
ggtggggaca	ccgtgaccct	tgcccagcat	tcccagccct	gctccataca	atagctccag	30720
gagacgcagc	agaaaagccc	caaggtaaaa	caaacagaaa	aatcaatgtg	ggaaactgta	30780
ctctgcccc	tgcttacaca	gtcacagtgc	ccttttagctt	caaaaaggct	cccagacacc	30840
cctcagagag	acattttggt	aattttgttt	aattccaggt	ttcccaagtt	tgttacgtaa	30900
cacctctgaa	aaacacatgg	aataggtgct	taagaaacac	tgatcttggc	tgggcgcagt	30960
ggctcatgcc	tgtaatccca	gcactttggg	aagccgaagc	tgggtgggaag	cttgagggtca	31020
ggagttcaag	accagcctgg	acaacatggt	gaaacccccat	ctccacaaaa	aatacaaaaa	31080
ttagctaggg	atggtggcat	gcgcctgtaa	tcccacctac	tccagaggct	gaggcaggag	31140
aatcgcttga	acctgggagg	tggaggttgc	agtgcagcca	gatcgaccca	ctgcacttta	31200
gcctgggtga	cagagcgaga	ctatgtcccc	accccccaaa	aaaaaagaaa	agaaaagaaa	31260
gaaacagtga	tcttgtccaa	cccatttgag	atgagacaa	tgagaccag	ggaggaaaag	31320
tgtactcaag	ttcacagagc	acattaatgg	ctttctcccc	attgtcgttg	tcccagccct	31380
aacccaaggg	tgtgaccatg	gctgtgtccc	ggtaataagg	agtgcctctt	aaccctctcg	31440
gttgacgtcc	cagcccagtt	tctgccta	caggacaaat	cacatcctgg	gaggtgaggg	31500
tggaaataag	ggagggaact	gagccagggc	agacagctct	cagaggaggt	ggctctgacg	31560
cagagcaggg	tcagaaccca	caccaggaga	gaatttaatt	gatcatgtgt	tccactcacc	31620

tgctcagcc	aagccctcag	ggcaggggaa	ggcaaagtca	ggatgccctt	cgcacacacc	31680
ctctcttggc	cccaccatcc	tccccaagtc	actagatccc	acagctgaga	aggaccttag	31740
gatccgtaca	aagcctaaac	acactccaca	gagggggaaa	ctgagactct	gaagggaggc	31800
ctcaacagct	ctggtaaaaa	aggcgtttag	gccggggcgca	gtggctcaca	cctgtaatcc	31860
cagcactttg	ggagggccgag	gcgggtggat	tgctcagct	caggagtctc	cgaccagcct	31920
gagcaacacg	gtgaaacccc	gtctccacta	aaatacgaaa	aaattagccg	ggcgtggagg	31980
cgtgcacctg	tagtcccagc	tactcgggag	gctgaggcag	gagaattgct	tgaacctagg	32040
aggcagaggt	tgagtgagc	cgagatcgcg	ccactgcact	ccagcctggg	cgacactgcg	32100
agactccgtc	tcaacaacaa	aaaaaaaaaa	atggtgttta	aacacatata	actaaattat	32160
ccttccccct	tccctgaag	tggtggctc	aggaaaaacc	tctaccact	caggcagagg	32220
ttttcctgca	ccctgcatcc	gtgaggcacc	actgccaaag	acgccaggga	aggctgcag	32280
gcctggagag	gggcaggggc	ccctccctc	caagggggcca	caaacgctgt	ctgcgcccag	32340
tacctgggt	aaggcgaggc	cgccgggcta	acccggggct	ggcggccttg	cagcgtgctg	32400
ggcaacagca	gctggggccc	caagactcag	cacgggaagt	cctcgtccaa	gtctggggcca	32460
agagcagcgg	cccagggggc	ggggccggcc	agagggagcg	gggagaggct	gagggggcgt	32520
gccagcgccg	gaccctgcca	ttggctggag	attacaggag	gcggggacat	agcagggagg	32580
agccgttgga	caagccccc	ccggccggca	gggaggtct	gaggtcaaga	gccggagaga	32640
agggatttag	ggccctgggc	caagttgcac	agcagggaga	aggggctgcg	cagaggggag	32700
gggagaaagg	gatccgcttc	cttcccttag	agctgtgaaa	tgtccccggg	tggaaattaa	32760
ggcggctgct	ggggagaggt	gaaattcagc	caaaaccacc	cagtacggca	gcccttctca	32820
gagataaaca	gtccgagcca	gcccggccag	gaacctcccc	ctccaacctc	cctaagcctt	32880
taacactcct	aagcctttaa	cgcgtttaca	cactcacata	aataaacaca	ctttgagcaa	32940
cacacataca	ccactcacca	catgtaatat	gtcaagccat	gtgcacgacg	aggtgtcgac	33000
aatttcatat	ggttcaacct	agtacactca	caaacacacc	taccaactca	tggctttcac	33060
agggacgggg	tcacacaccc	actctccac	gacatggcaa	gcgtgcacac	gctatctcaa	33120
gctgctccct	ccccctcaag	atcatgttac	ccagttttat	tttcttccca	gcacctatga	33180
cgactgacat	aatttattag	tttacttgtt	tattgggtta	tctgtgcccc	tcacccccaa	33240
aatgtaacct	ccagcagggg	ggatgactcg	gtcagtcctg	attgtgctgt	agtccaggac	33300
ctagaacaga	ctccatgga	cattcatggg	ctctgtacac	acaaacacac	acattaacat	33360
acaccccgac	acacagcctc	atccacacac	acacagcctc	acacctgctc	tttgacgcca	33420
cctgcacagt	ttctcacaca	ctcacttgat	ctagtgtat	gcgtccacag	gcccctcccc	33480
cagcccactc	atactgccct	caccccactc	actctgcctt	caccccactc	gggggaactc	33540
tgtgtccagg	ccaggcctgt	gacactcacc	gtgagtgaag	tgaacgttag	gcagatgcca	33600
ccaaagccat	tcagggacag	cgccagggaat	atcaacggag	acagagctgg	aaaggggaaa	33660
gcagcagatg	agggcatttg	gggagctgtg	ggaagccaag	ggcgggagct	ggggtaaaca	33720
tccgccttca	tcccacctat	tcttttcttg	tgggggcaca	agaggacaga	caactcacct	33780
tccacgtccc	gggagggcag	ggccatgagg	gtgcaggacg	cagtgaagca	ggcactgtgg	33840
agacacaggg	aagggcgagg	ggttggcctg	tgagcaccct	ccctccctc	cccctgcagc	33900
acggtccctg	tcctcccggt	ccccatagcc	cagccacctc	acctgccaac	cagccgcacg	33960
ggtcgggggc	caaagcggtc	catgaggatc	cccagtgcca	gggtggtggc	gctgagcacg	34020
aaggaaacca	tggtgaagcc	caggttgagc	atctcgtcct	gctggtcaca	gcctggccac	34080
ctgcgtgct	catcctgggt	ggtgttgggt	ctgctctcag	ctgaggaggg	ggaagggagg	34140
gctcagcaca	tgacaccagg	aacagctggg	cacaggagac	agcagccac	agtcaggcgg	34200
cctgctttca	aatcccatgc	caagtgcctt	tgggggtacc	ctagagtcac	atctcctctg	34260
atggggctgc	tccagaaatg	gcagccatta	gtacctgacc	ctgggagagt	cttgtgcaca	34320
cacagcctga	ggcttcaact	agctcaaatg	aaatactgga	cataaaagta	tttactaagt	34380
tgtaatatgc	actcagtgtc	caagcttagg	gggttgtgga	cccccaacaa	gaagtgcctc	34440
catatctaga	ggcaaaaggca	aaggcagtga	gtggtactct	aatggctata	acaagaattc	34500
attaaaatgg	ccggcggtgg	tagttcatgc	ctgtaatccc	accactgtgg	gaggctgaga	34560
caggcagatc	gcttaagcct	acaagtttga	gaccagcctg	ggcaacatgg	taaaacccca	34620
tctctaaata	aaaaaaaagaa	atthagaaag	aacactaaaa	cttagaggaa	gctttcccgga	34680
taaatgatag	tctgataaaa	taatagctaa	tacttattga	gcacttaact	atgctccagg	34740
cactgttata	agtcagttaa	taaagtatcc	cgttccctag	gtgatgaagc	tgaggcacag	34800
aatgagaacc	aggcactgcc	ctccagctccc	ctctagaagt	ccacttgagg	gacttgtcct	34860
taacggtaaa	ctgccaaact	ggagttgtga	caagttaaag	agaaaagcta	gtgataggag	34920
acaaagggct	gcttcgcttt	actcaatgct	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34980
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35040
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35340
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35400
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35460

gaaaggaaaa	tgcaaagggg	gaaggggtca	cacttgggga	aggtttcaga	caataccgag	35520
tggaaagggg	gatgccaggt	gtggggagta	acagatagag	gaggcaaagt	gagtggagac	35580
caagccagac	cggggaggag	ggggccacag	ccaaggtgag	acaggtcagc	agccagaaac	35640
cgaagcagac	acttgcaggg	tgacccccgc	cctctcttcg	tggcaatctg	agaccgagga	35700
cgtggagacc	ctggagagcc	cccaaccttg	tttctggggg	gtgggtcaga	gaggaagcct	35760
ctcatccccc	ggcaccagcg	gccttcccgg	gaggtctaac	acgcagatac	ctggataggc	35820
gtccatcact	ccccgggcca	gagccacca	acgtctctcg	aggctccgacc	ttgtccctcc	35880
ttctacccca	cagtccccag	tcctagctca	tctgcataaa	gctccaatta	acatgttttt	35940
cctttgctat	ttgcgatccc	agaactcggt	ccccaccccc	agcccgtttc	ccgccgcttc	36000
ctcgcacctg	ggagggcggc	cccattaacc	ctcgcgaccc	gggccgctcc	tggcggctcc	36060
gaaccccgcca	ccccgtcccg	cggcgggggt	ctgggggtga	ggggcgcgcc	ctggggcaga	36120
ggattgcgcg	gcagggctcg	ccacagggca	gagggccagg	ctctccggga	aaaaggcagg	36180
cgcatatatg	cccccttttc	tgggaaaaga	cggggagggg	ggcttctcct	gggagactcc	36240
aggcttcgaa	attcctcggt	ccctatcctc	cgccccccgc	accctcctc	ctccccgcca	36300
cgcacctctc	ccccccccca	gccatctggt	ccactccgca	gcgcgcgcgac	aaacacggct	36360
ccagctcgct	tccgccccctg	cccagcccc	tccccaagcc	ccggggagtg	ggggagtgag	36420
cagacgcctc	tctcctagga	ggcgggaatt	tctgcctcca	tctcccaccg	gggtccggct	36480
ggccagaggg	aagcttgcag	acccccacc	aaccaccacc	accgttgcca	gggccggtga	36540
ggctgcagat	aacgttgca	aggacgggag	tcggggaggg	tgtaggcgca	gtttaaagga	36600
cgggcagagc	aagccccggg	aagaggcagg	ggttttccct	cccgggtcgc	cgccccccgc	36660
accctcggag	ccagccgcag	ccacgcagcg	ccgcctgccg	ggcacaccaa	ggacctggcg	36720
cgcacgtggc	gcttaccccc	acccccgggt	ccgctcctgg	ctcgcctca	gcctccccag	36780
actattcgca	aattgaggat	cccggaacaca	gagtgccagag	accccgcaa	gcctactgaa	36840
agccagccga	acccgctggt	gggtgctagc	caattctgat	tttgtacttt	acaaaaacaa	36900
aaaaagtcag	tgttggaagt	cgggagtcgt	ggctcagagc	agcagggatc	tgcgatgtga	36960
ctttgccaa	tctccagacc	cctgaggaca	ggttttcccta	tctgaaaacg	gaggggacag	37020
tctctcttat	taacttctca	agagaaacaa	agacaaaggg	agggaaaatg	gcttagctgg	37080
aatgctgtct	tacagagcca	acctttggag	gtgggggaga	tggccaaggc	ctctgaggtc	37140
actcttgccc	ccaggagcag	ctgagaaccg	gaaagaagct	tgggacctcc	tttctgcaga	37200
gctatccttt	ccacagactg	ccgaggttcc	aaattgagct	ccaccacctc	acactgtgtg	37260
cccttgggtg	tgtgccttaa	cctctctggg	cttgtttccct	acagcgacaa	gaaagaatga	37320
caacaccaac	ctcttaggct	atagtttgga	taaaatgaga	tagctgtgta	gaacagacag	37380
atcctaaacc	aatgttagtt	ttcccttcat	ttggggactt	gctctaacct	ccagggtcta	37440
tgtcccagag	gcacaagcag	gtgcagggtc	ggataaataa	ggtatgtctt	tctgcaggat	37500
ctcttgcctc	cactgatggt	gtcttctcct	gatataagata	attttaaagc	ttcacgttat	37560
ttattttatt	actttaaagc	ctcactttaa	tgttaaaggt	aaatgtaaat	atagtataac	37620
aaggaagctc	aaaatttgca	taaagtttta	agataaaaata	ggagactcca	aaaaagtggt	37680
actttcggca	ggccctaggg	atgctatggt	gggaagtttg	agtcataacct	tagcattctt	37740
tctaaagcat	tctgtcctaa	tcctctgtat	ggagaaaagc	cagcttcctg	gatgtacccc	37800
aaatcttgcc	aagttagggg	caggagctgg	actccctcca	agcactaagg	gcagggcctg	37860
ggtgggaaca	gggaggtgag	ccagacagcc	agaggcgaa	gggctggcat	gccaaagctc	37920
ctagttaatg	cccagctgag	cctgggtgaa	gaaggatggg	ggtgtgggga	agacaccccc	37980
caccaaccgc	caaagacagg	cgcacaccag	ccagtctctc	acttcccttt	ttatttccctc	38040
taagacttgc	aagcagcagc	accagagagg	gaacctgccc	tcttggccct	ggaaggggccc	38100
gagccccaac	ccctaaccga	ggacacagct	ggcacctcag	gcccccttcc	ttctgaaagg	38160
agggctgtgt	ctctctcaca	ttcacacata	cacagacaca	tgcattgtgtg	cacactcatg	38220
gcacatggga	cctcaggggt	agcctgtttg	ccgatccccc	caagaggtac	caggaggcag	38280
accgctagaa	ggagataaga	ggcaccctgg	tctcctccaa	cccaaggagg	aagaaagctc	38340
aacccctcta	ggatagggac	tgtcttcagt	caatggagcg	ttgacttagg	gggcgttttt	38400
gaaggttttt	tttctcctt	tttgagctg	ttacaaaaat	agaacttctc	ttgggtattta	38460
taaatctacg	gccatggctc	tatgtgcatg	ttacaggtag	aaaagccata	tgggggactc	38520
cttttggttg	ctcaggcctt	gattgcctgt	catccaggtc	ccttgggtctg	agaagtctat	38580
gcggtcacct	cagagccgct	aagcaccttc	agtgggcccc	tccattggc	ggcgactacc	38640
tgctggagcc	gggcacggta	atagaagagg	taggaaggca	acaggaatcc	caggagttag	38700
aatagcagga	ggcccagatt	caccttttag	gcaaggagag	agaaacagag	tcaagttagt	38760
agtcacttgc	ccttagcctc	ccacagggag	gagaaggcgg	ccatttttct	ccaggtcctg	38820
agccagaata	aatacagcta	gtacttatta	tgtgtagtca	ttgttccacc	agtatctcac	38880
ttaatgttca	gcaattctgc	aaagtggctg	agatgagact	tctcaggtat	aacaagtggc	38940
agggcctggt	gggtgcccac	accatatggc	actcactagg	taggtatgag	gaaggcacag	39000
cactgtagga	gtctgggctg	gtcaggctgc	tcccgaatg	gggccttctg	ggctcacccc	39060
tctgaccttt	ggagatgtta	accaatggga	tcccggttcag	ggtggcgaga	ggaggctctc	39120
agacacagtt	caaggaactg	ggatgcacag	cctggtggac	agaaggcttg	gaaggcccag	39180
gacacgcggg	ctctgactcg	gttcacatcc	cactctgcat	tactcactgt	gtgacttttg	39240
gcaataaatg	gcaattctta	ctgagtgcc	ccttctcagg	gctgtgtgtg	cgaagatgta	39300

agttaaaaaa	aagtatgcat	catgcttagc	acatagttag	tgcttggtta	atagaagcag	39360
ttattttcatc	acaattcttt	gggaggaggg	tttacgtgtg	ggtggcccca	cagggcagat	39420
gaaagatcag	cgctcagggag	gcagatgagt	tcaatgtaag	gaaaagactt	actaacagca	39480
gcagggctgc	ctcgtgcagg	agtgggtgcc	ctaccactga	gggtatctaa	gctaagaggg	39540
aagggtcccc	tttcaggggt	gctggagaca	ggatcccaca	ctaggtagaa	ctggattgga	39600
ccaatgggtgc	ctgaacacag	gccaagagt	caggactggc	cacttcacaa	agcacctgga	39660
gtttactaaa	aacagactcc	taggaggtca	ggcactgtgg	ctcacgcctg	taaccccagc	39720
actctgggag	gccaaggtga	gaagatcatt	tgaggccagg	agtttaagac	tagcctgtgc	39780
aacatggcaa	gaccctgttt	atctgtacaa	aatttttttt	taaaaaatta	gccagggtatg	39840
gtagccatca	cctgtgggtg	cagctactca	gaaggctggg	gccggaggat	cgcttgagcc	39900
caggaatcag	aggctgcagt	gagctgtgat	tttaccaccg	cactccagac	tgggcaacag	39960
aacaagacac	cttctctaca	aaaaaaaaaa	aacaataggg	ccgggcgcgg	tggctaaggc	40020
atgtaatccc	agcacttttg	gaggctgagg	agggcagatc	acgaggctcg	gagatcgagg	40080
ccatcctggc	tagcacgggtg	aaaccccgtc	tctactaaaa	atccaaaaaa	aaaaaaaaaa	40140
ttagctgggc	gtgggtgggtg	gcgcctgtgg	tcccagctac	ttgagaggct	gaggcaggag	40200
aatggcatga	acccgggagg	cgagccttgc	agtgagccga	gacgcacca	ctgcactcca	40260
gcctgggcaa	cagaatgaga	ctccgtctca	aaaaataaaa	ataaaaaata	ataataaat	40320
aaaataacaa	taaattaaaa	acaaaaacag	actcctacgg	tcaggctgag	atatcctgat	40380
tcaggggact	ggggaatctg	tatttttaac	actccgtgag	gggttctaaa	aggcagacaa	40440
cttggaacc	tgcagattag	agacctctga	ggtgcctctg	gctgagatga	gtgagggatg	40500
gcaccacata	caaggcccta	cccctgcccc	caggagagt	gctcctgctc	ccccacacc	40560
aacctcgcgt	ctaccccaga	agggctctcc	tttcaggggt	cccaccatcc	ccatgaaaag	40620
tggctgctga	agcaaggcga	acacagcact	ggtgagggac	tgcaggcctg	tcagcgtccc	40680
aaaaggggtt	ggatgggaac	ctgtcccca	aacgggagat	caaaggggtg	tgggggcctt	40740
tcagcccagg	caagaacttt	ttcttttct	tccaacatg	ggnnnnnnnn	nnnnnnnnnn	40800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	40860
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	40920
aagaaaaaaa	aatcttaaa	aataaggata	taagaaaaga	aaatatTTTT	gtgtagctgt	40980
tcaatgtttg	tatttcaagc	caagtgttat	tacaaaacag	tcaaaagtTT	ttaaaaattt	41040
aaaagtTTT	aaagtaaaaa	agctaagtaa	gctagggTTA	atttttttat	cgaacaaaga	41100
aaaatatctt	tgtataaaact	tagtgtagtc	taagtgtaca	ttgtttttat	tttattttatt	41160
ttttattttt	ttgaaatgga	gtttcactct	tggtgccag	gctggagtgc	aatggcatga	41220
tcttggtcca	cggaagctc	tgctcctgg	gttcaagcga	ttctcctgcc	tcagcctccc	41280
aagttagctg	gattataggc	acccgccacc	atgcatggct	agtttctttg	catttttttt	41340
ttgaaatgga	attttgctct	ttgacccagg	ctggagtga	atggtgcaat	ctgggctaaa	41400
tgcaacctcc	acctcccagg	ttcaagagat	tctcctgcct	cagcctcctg	agtagctggg	41460
attacaggca	tgcaaccacca	cactcggcta	atttttgtat	ttttagtaga	gacagggttc	41520
tcaactaaag	agaacatgt	tgccaggct	ggtctagaat	tcctgacctc	aggtgatcca	41580
cccactcgg	tctcccaaag	tgctgggatt	gcaggcatga	gccaccatgc	ccagccagta	41640
tacagtgttt	ataaagcctc	cagtagtgta	cagcaatgtc	ctagaccttc	acattcactt	41700
actactcact	cactcactca	cccagagcaa	ctgccagtc	tgcaagctgc	atgcatgata	41760
agtgccttat	ataggtgaac	cattttttta	tattttatac	tatatTTTTA	ctgcaccttt	41820
tctatgatta	gtacacaaaa	tgcttaccat	tggtttacaa	ctgcctacag	taatcagtac	41880
agtactatgt	atgggtttgt	agcctaggct	ataccatggt	gcctacgtgt	gtagtctct	41940
atactgtcta	gtttgtacac	tctatcatgt	ttgcataaag	ataaaatcac	ctaatagcac	42000
atttctctga	gtgtattcct	gttggttaagc	aacacatgta	taaacattta	caagaaatag	42060
ctcaaatTTT	tttttctttt	gatacagggt	cttgctttgt	caccagggt	ggagtgcagt	42120
ggcgcaatct	cggcgcaactg	cgacatctac	ctccccgggt	caatcgattc	tccggcctta	42180
gcctcctgag	tagttaggac	tacaggcacg	caccaccacg	cctggctaatt	tttttgtat	42240
ttttattaag	agatgggggt	ttgccatggt	gctagggctg	gtctcgaact	cctgacctca	42300
ggtgatctgc	ccgccttggc	ctcccaacat	gctgggatta	caggcatgag	ccaccatgcc	42360
cagccattac	gttttttttg	ttgtttaatt	tttttttttt	taagagacag	attctcactc	42420
tgtcatcaag	gctggagtgc	aatggcacia	ccatagctca	ctgcagcctc	caactcctgg	42480
gctcaaggga	ccctcctgcc	tcagccttcc	cagtaactga	gactacaggt	gtgagccacc	42540
atgctcagct	aattattttt	tatctttttt	ttttttaga	gggggggtct	ttctatgttg	42600
ctcaggtttg	tctcaaaactc	ctgggctcaa	tcaattctcc	tgctttggcc	tcccaagggt	42660
ctgggattac	aggtgtgagc	ctgaaaacct	tctagtgtgg	aagtgaaga	tagggccagg	42720
ccacttatgt	tttcaagtta	agcaagggtt	aggtcactta	tgaagcctga	ctagttttgt	42780
ttgcttaagg	gatctgcagg	cctgacctcg	gttttcatTT	gttttaacag	tgtctatgtg	42840
tatgtgtgtg	tttatgtacg	tgcatgatgg	ggggaaagct	cagaaatcaa	gtaagccaaa	42900
cacaaacatg	taattataag	cagggataaa	ttctatgatg	aagaagtatg	ggccacggga	42960
gagtacttgt	gccagtctgg	tgatcaggaa	caatgtcctt	tgggaagtga	catttgagcc	43020
atgccctgaa	gtacggtagg	agttgggttag	gggtgaggca	gtaagaccca	gagctggggc	43080
ttcctgcaca	agctcagctg	ggcactgagg	acccagtgga	ctctgctaca	gggcagttag	43140

gagcagaaag	gctgaggaag	gctgggtgtg	gtggctcaca	cttghtaatcc	cagggccttg	43200
agaggctgat	gggggaaaaat	cggtagagct	caggagtttg	agaccagcct	gagcaacata	43260
gcaagactcc	atccctgtaa	aaagctttta	aaaatttagct	gggtgtgggtg	gtatgcatct	43320
gcagtctcag	ctactcaaga	ggctggggta	aggattgctt	gagcctagga	ggtggacgct	43380
gcagtgcgcc	acgattgtgc	cactgtactc	caacctagga	gacaaagcga	gacacctgtct	43440
caaaactgaa	tgaataggct	gtgtgcgggtg	gctcactcct	gtaatcccag	cacttttgga	43500
ggctgaggtg	ggtggatcac	ctgtgattgg	gagtttgaga	ccagcctggc	caatatgggtg	43560
aaacccgata	caaaaattaa	ctgggcatgg	tggctcacat	ctgtaattcc	agctactcgg	43620
gaggctgagg	catgagaatg	tcttgaaccc	ggggggcaga	gggtgcagtg	agctgagatc	43680
gcaccactgc	actccagcct	gggagacagc	gagactccat	ctcaaaaaaa	aaataataat	43740
aataacaatt	aaaaaaaaat	taaaaggcca	gggagcactg	gcagcctgtc	caaggtttca	43800
ggctcacttta	gtaaaggagg	aacaatggct	cctcccagga	cctctgggat	ctcagcattg	43860
atacgacagt	catggaaatg	ctagggccca	ggcagaccat	ctcagggaaa	acaagtggct	43920
ctgccctgcc	ttggccactt	cctggccctc	tgcattgcccc	agggtctcag	caccaagctg	43980
ttctcagtga	gtagctctca	tttagtgcca	gggtctctcg	gcttacatcc	tacgatgacg	44040
atggaatgca	taaaagatgg	ggctgtgata	gccagagct	aggggtttga	atctcatgag	44100
atgttcatgg	agccctggga	gggagctcag	tgcaagttca	tttctctttt	ttggttgaga	44160
tggggctcag	aggaggaagg	acttgttcaa	agacacacag	ggagtgtttc	agtgtgggac	44220
ggaggtttat	ggagaaaggg	tgaccatcca	aggcttgga	aaagatcatg	acttcgacca	44280
gcaagcctca	actctgtaga	cttggtgggg	gccaggccct	cccaaacaca	cctgacaggt	44340
gtctgtgggtc	ttggggacat	tgtcgtctcc	cttctctctg	atgctctgct	gtccctctcc	44400
catgaagcgt	atctcttcgc	cgtcccccct	ccttctctgag	agaggatggg	ttctctctctg	44460
accaatactg	aagatcttta	gtaaagttct	cttttttttc	attttctgaa	agtccctctc	44520
ttgagaaatc	aggacaagtg	agtcagggcc	aggacaaaaa	acagtgtggg	acgagtgtgg	44580
tggtctcagc	ctgtaatccc	agcacttttg	gaggccaagg	tggcggatca	cttgagggtca	44640
tgagtttgag	actagcctgg	ccaacatggt	gaaacctcgt	ctctacaaaa	tacaaaaaatt	44700
agccaggcgt	ggtggtgcat	gcctgtaatc	ccagctattc	gggaggctga	ggcaggagaa	44760
tcacatgaac	ccaggaggcg	gaggttgcat	cgagctgaaa	ttggggccact	gcactctggc	44820
ctcttgccaa	cagagccaga	ctacctctca	aaacaaaaac	aaaaacaaac	gacaaacagt	44880
gtagactttg	tgtttttctc	aaaagcactg	tcaagccagt	gcccgagca	gtgggcctag	44940
acacctccag	tcttgccctca	gggtcagttt	ccagcctccc	tggacacttc	ccccaggat	45000
gtgtactttt	tgattgtcct	aaatccagag	tctgtggcct	gacctgggtt	gtcacagctc	45060
tcagtccctc	cccatcccga	atcccaggga	gccgcaggtg	tgtgcagaag	aggcacacca	45120
caactcaatc	atcttgcatc	ctcgtgggac	ccaatccatt	ggcttgggtg	tgtacagact	45180
gagcctcatt	atagccgttc	gttctctgtg	acctttccag	atcaatctgc	cagcttggct	45240
tctccgagtt	tccgttgcca	gcatttctcc	aatcccatca	tgtacttttg	acctctttgt	45300
tgggtggcct	gctttatctg	aaattttcag	atttgacttc	aggtctctcc	tttgtccctt	45360
aatatggctt	aatgggtggac	cctgtcaggg	gtagacaaaa	tattgaggag	ccctgacttt	45420
gaggtgcaca	agtttagagg	ttagacaagt	ccagccacaa	ccagcccaag	ctgcagtgtg	45480
gggaggcctg	tccagctgct	ccacggttga	gggtggagca	tacaggaagg	cttccctctt	45540
gctgcagccc	aggtgttctg	gctgccctag	ctgcctggct	ttggtagaag	aaagaaaggc	45600
tctgtctctg	acttgtcaac	taatggcact	atgagattgc	acataattaa	cctgggtctg	45660
ctcttccaaa	agccttgggc	ctctgactgc	aacatggagt	ctgggtatca	ctcccatcc	45720
ctgcgcctc	cactgctctc	ggcgctaggc	gtgtgcctaa	tcacttaatt	tctctgtgct	45780
gectcttagg	tatcacttcc	cctgatccca	aatacttacc	agggtgtggg	tgacacctga	45840
ctagttactc	cttgagggtg	tctgcttctc	accggggact	ccgaaaccaa	acgaaaagca	45900
aggccaagcc	cagcctaaag	gacgcttctc	acatgacttc	aggcttgcgg	gggctggagc	45960
gtgggggtgg	caatggagtt	ggggggggct	cagggagggg	atgtggaagt	gctttgcttt	46020
gcaaaactcta	gagaaccgtg	taaaataggag	tgattattct	gtcccttccc	tttctttcca	46080
acagggaatca	gcatcccaca	gcccattgtc	agctatgaag	aatggaaact	gaggctccgg	46140
gaggggtata	gggaggagcc	agcagggtct	tgagttcata	ttagtgcctt	ttcctccata	46200
ggcacatctg	tgttttcttt	tattttattt	tgaatttaat	tttttttttt	tttggcagag	46260
tcttgctctg	tgcgccaggc	tggagtgcag	tggcgcggtc	tcagttcact	gcaatctccg	46320
cctcctgggt	tcaagtgtat	ctcctgcctc	agcctcccga	gtagctggga	ttacaggtgt	46380
acaccaccac	acccagctga	tttttgcaat	tttagtagag	acagggtttc	acagtgttgg	46440
ccaggccttg	cttgaaaatc	tgacctcaag	tgatctgcta	gcctcgccct	cccaaagtgc	46500
tggtattata	ggtgtgagcc	actgcgctcg	gccacatctg	tgttttaaat	gagaggaaa	46560
gggataatgt	gcattttgtg	gaagcttggg	ccgtttgtgt	ctaggactct	tatgatcttc	46620
ataagttttc	ccccagggag	gacactgttc	cacttaggga	gtcaggaccc	ccagtcctta	46680
caagattcag	cctctcaaaa	tggagacagc	agttccaggc	ctgggctggg	ttctgttcac	46740
actaggagag	ggcaagtgtg	tggtgtttgg	gatgtgggga	agtattatga	aaacagagat	46800
gctccaattc	ctagtgtatg	gaaaccatta	agctacttgg	catcttaaaa	ccaagagcgg	46860
ttcaagttct	gagattgtta	acacacctta	caacaccgcc	gccgttatta	ggaagaagct	46920
ctgtttgatg	acgtcccaca	ctgtgggtac	ctttatgaac	aggaatttgc	tttttcaaat	46980

cccagagaag	taagattaaa	gttggctggt	ctccatcctt	gaaaaatttg	gttttagggg	47040
gaattcaaga	atgactgacc	atacagaatg	gggagcaaac	ttgggaagaa	agaaggcaca	47100
gttcagagct	ctcccaatag	tcacccctga	actgcacccg	gaccatcagt	tatctctgtg	47160
ggtagagctc	aggaatctaa	aatccatttt	aaaattaaag	tatatacggg	ctgggcgcgg	47220
tggctcatgc	ctgtaatccc	agcacttttg	gagggcgagg	tgggaggatc	acgaggctcag	47280
gagtttgaga	ccagcctggc	cacatggtga	aaccccgctc	ctactaacia	tacaaaaatt	47340
agccaggcat	ggtggcagac	acctgtagtc	ccagctattc	ggaaggctga	gtcagaagaa	47400
ttgcttgaac	ctgggaggca	gaggttgtag	taagccaaga	ttgtgccact	gcactccagc	47460
ctgggcaaca	gagggagact	ctgtctcaaa	aaaaaaaaaa	aaaaaattaa	agtatgtcat	47520
acatactgtt	acaggcacag	accttaagtg	tacagcccaa	tgaaatttta	cacatctata	47580
cagctatata	actaccacct	atatcaagac	acattccagg	aactcagact	ccatcatacc	47640
cctcctcagc	agaggtaaca	gacccacacc	tctcctgctc	cggtggtaat	taaccactat	47700
tctaactttt	ctatcaatta	gttttgccca	ttcttgagct	tcacacagat	atacattgtc	47760
aggcatgatg	actcatgcct	gtaatctcag	cactttggga	ggccgagacg	ggagtatcac	47820
ttgagcccag	gagttggaga	ctactctgga	caacatagtg	agacccccga	ctctacaaaa	47880
aaaaataaat	agctgggtcat	ggtggtgcgt	gcctgtagtc	ttagctattt	gagacgctga	47940
gagaggagaa	tctcttgagc	ctgggaggtt	gaggtgaaag	tgagccgtga	ttgcaccact	48000
gcactgcagc	ctaggtgaca	gagtgagatt	ctgcctcaaa	aaagaaaaaa	tatggccggg	48060
cgcggtggct	caagcctgta	atcccagcac	tttgggaggg	caagggcggg	ggatcacgag	48120
gtcaggagat	ggagaccatc	ctggctaaca	cggtgaaacc	ctgtctctac	taaaaaataca	48180
aaaaaagaaa	gaaaaaaaaa	ttagccaggc	atggtggcgg	gctcttgtag	tccagttac	48240
ttgggagggt	gaggcaagag	aatggtgtga	accggggagg	cagagcttgc	agtgagccga	48300
gatcgcacca	ttgcactcca	gcctgggcga	cagagtaaga	ctctgtctca	aaaaaaaaaa	48360
ggaaaaagaa	aaaatatata	tacattgtgt	actttttggc	atctgggtta	ttttgctcaa	48420
tatcacatct	gcgaaattaa	tctacactgt	gtgtatgaaa	ggttggttct	ttttgttgtg	48480
atgcagtatt	ccgtcgtgtg	actacgggac	aatttgctta	tccgtattcc	tatcggtggg	48540
cattttgggt	gttaccagggt	tctggctggt	atgaataaag	ttgctatgga	tattcttgta	48600
cactacttct	ggtgagcgta	tgcaactcatt	tcgcttatgt	aaatatcttg	ggtggaatta	48660
cctgatcata	aggtaggtgt	gttggccttg	taatgtgctg	acttggttat	gctgaattcc	48720
ctttttttgtg	tatttctggt	tagagcggaa	catgaggggtg	tctcttcagg	gaatctggag	48780
ggtggaaggg	aagcaggagt	cggtttcttg	ctcacacatg	ttgtgactga	actgctggta	48840
cacctgggtg	gcatggagct	ggcttctcct	ttggcggtgc	ctactggttg	ggcagggtg	48900
tatgtggtta	gctccatgca	atgaacccgg	gcttctgcaa	aatacattaa	caacgacaga	48960
gacaacaaaa	gctgatgtgg	atttaaaggc	ttcagttcan	nnnnnnnnnn	nnnnnnnnnn	49020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49260
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49320
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49380
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49620
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49680
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49740
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49860
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49920
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49980
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50040
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50340
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50400
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50460
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50520
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50640
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50700
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50760
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50820

gctaaacttt	ctccaggggg	taaagttcaa	gttgattagt	caattgatta	attaattcat	50880
tatgtaatgg	aaaaactcct	tctatgacct	gggcagagtt	ataggcagtg	aacaagacag	50940
acaaggtcct	tgttgtcatg	aagtttgctt	tctgaaggag	agagataata	aacaagaaac	51000
cagtaagaaa	gcaagattat	atcatttttg	taaagtgtct	tgtggaaata	aatgtgatga	51060
tgtgtaacaa	aagtaccaa	taggagagt	gggtgggtgg	gcttctttta	gaaagagttc	51120
tcggagaagg	cttatctgag	gaggtggcct	tttaaccagt	acaaatgctt	tagcttggcc	51180
agtggagctg	ggaccaggat	gacaagggtc	acttgtcatg	ccagtgaagt	tgagcttgta	51240
gacaagagcc	tgatcatgaa	agactttgca	gatgggtgta	atgggttttg	gttaattgct	51300
actatgtggg	aagactttga	atgggaagca	tggggacaat	ggcctgtgat	acatgttate	51360
aaatatggtc	gcaggggcta	gtgaggtggc	agcagagata	gggagaagta	gacggactgg	51420
ggaaggtaga	agatggggca	ggggaggcaa	ttactgcaaa	gacataattc	ttctaagctc	51480
actgagtgtt	catggtctct	gggagcagag	gttctctggg	gggaaagagg	ataatgtcac	51540
ttcctgagga	agcgggaaga	acccatctga	gacgtgggga	ctgtgctggt	tcgtttctaa	51600
ggggccttcc	agatctcaca	tgccaatcgt	cttgggtctat	gtcaattggt	ggggcatcca	51660
aatggggaac	tgttgtccag	gccgatttca	cagaacaacc	gcccagtcga	tatctcccga	51720
gccattcacc	cttgacagtgg	cgttagctct	ttcaccagct	tttatctgcc	ccgtgggat	51780
ggtggccaag	cccagttaac	aagcagttga	tcagccccag	agatcaggtc	cctggagtct	51840
gtcacttttc	tgaggggtgg	gagagaatcc	tggagcagaa	catgtaacta	gaagggccac	51900
ctggcttctc	atgggtctgag	ggagagaatg	gtgggatctc	tggcctgaat	caaacctccc	51960
tttctcagtg	tccatctttac	ctctctgctg	taccttcggt	attttccagc	agctcctcag	52020
cccgttctctg	tgggaccctt	ctctgccaat	ccctacaccc	actgtaaat	tcaccgtggg	52080
agggagatgg	gccttgaggg	ctgtattagt	cttctattct	gcataacaaa	ttgcctcaaa	52140
tttagcagct	tcaacaact	catgtttatt	agctcatcgt	gagttcatca	gcagtgtggg	52200
cccagcatgg	ctaggttttc	tgtctcagggt	ctcacaaggc	taaaatcaag	atgttgtctg	52260
ggctgtgtgc	tcatctggag	tttaggggtc	tcttccaggc	tcacgtggtt	gtggcagaat	52320
tctgttccct	ggagttgcag	ggctgaggtc	ctgttttctt	gctgactgtc	agatgagggc	52380
tgctctcagg	tccctcaggc	tgcccacatt	gcttgccacg	tgcgtggtct	ttccatcct	52440
tgaagccagt	gatggagaat	ttcccttgga	ttgaatcac	cacatgggtg	gactctctga	52500
cttcagggaag	agagccctgt	ctcttttatg	ggatcacctg	attagatcat	acccatagag	52560
ggcagttcct	tttccctaaa	gtcaactgtg	gcatgtaaca	tcacacaacc	acaggagtaa	52620
aatccatcat	atttacagtc	ccagggatta	tgcacagtgc	accaggggac	aactgaattc	52680
tgctgtctaa	aagggccaag	caggacttta	ttggtgaaga	acagtgggaat	gtcattcttg	52740
gttcttccag	aaaaaaatca	ctcagtaaa	ttagagggtc	tcttgccctt	tgggaagtga	52800
tcaaagaatc	tcatggagg	tttggacctt	caccctagaa	acatcacacc	atgttttcta	52860
taattgcagg	gttcatggtc	ccttgaagcc	tattcatagt	ttccagggtg	aaaagctctg	52920
ctgcagggtg	tggggaggga	tgaggtgga	ggtgagggtc	gaatagtgtg	agctgcatat	52980
ctggagctgt	ggtggttttt	ttagtcttta	agctgtcatg	tgttgggggt	tgggcatggg	53040
aggggcatcc	caagagctcc	ttggtattga	caccatctcc	aaggtgatct	ctgctctgcc	53100
tggtgcacac	atgtttttct	cctgttgcaa	cagccactc	ttgtagaaga	gcagaccct	53160
cagtaccagg	tctgacctg	gacagcttgt	accaggagct	acagcacact	ccccacaa	53220
cctaaagtgt	ggatgagccc	cccagagaat	agatcagaaa	agattaaatg	cagagggtgat	53280
ctgtcaggtc	ccctttggaa	gtgctggtat	ggagaggatt	gactgagctc	gttttaggaac	53340
ctccaagctc	tgtagtaact	ttagggttag	aaaggaggat	gcctaagatt	caggatcctg	53400
cagtgatgag	tcaacatttc	ttgggggaag	aggcagggct	gaggattaaa	cggagatgat	53460
gggtatcggt	ctcttgctca	aaggcactgg	accccaaggc	ctccagctct	tcgctcccat	53520
ttgaaattca	agtcctgagc	acaccacagt	tgtgatgcag	ggaaagaatg	tgcttatcag	53580
agagcctggg	caagtgggccc	ccttgtgagt	accgttcaac	ctcatttatg	tcattggcac	53640
caaaagtaga	catcagctctc	ttgaaagttt	gattaatgct	ggtcacactc	aaagaccctg	53700
ggtagcattc	atttactaag	caattactaa	ataccagttt	ctgtgctaaa	tgctgcatca	53760
gtcagggtct	ttaatggcag	gcagcagaaa	ctctccttgg	ctgatctaag	tagaaaaatc	53820
caggactgaa	aggaaacgga	gtagctcatg	aaattgcagg	aagggccgga	aaaccagaca	53880
tggagccaaa	gtcagggtgc	agaacaggtc	tagggaggat	cccactgctg	ctgagacctc	53940
gaccttgtgt	ctggcaccca	ggatgttgta	gggctcagac	cctggatcaa	tgatcctg	54000
agtgcctctg	tgggtactgc	aactccagga	actcaatctt	gtcaacgcca	ccgcccagaga	54060
gaggccttct	tggcctccat	ctttttggtc	actagctcca	gattcaaaat	cttgaataga	54120
tgcttcttct	ctttgataga	gcccagctcat	atgcgttagc	tgcaaaggaa	gctgaaaatc	54180
tattaggaac	ttttgtcttc	aaaaatgaga	ggcctgtcct	ccaccaagat	ccataggaaa	54240
tggaatccaa	gaaaccacag	gaaggggtga	ggtgactggg	cagctcacag	catgcatgct	54300
acatgtgaat	tatctcattc	atttctcaca	ctacccagtg	aggtagggtat	tgtcatccct	54360
acttcataaa	tgatgatatg	aggtacagaa	agtttaagga	acttgcccag	gacacgcac	54420
gcagctatta	agtgctagac	ccagtcaatt	ttggtctgac	ttggactgtc	tgactccaga	54480
agccaccctc	tcagacactg	ctgtatactt	ccagtgaatg	ttgatgaaat	tttcagggtt	54540
gctaagctgt	ggatttcaga	tcctggattg	tatgacctaa	aagagagact	tccctaggag	54600
tgagggtccc	tgaacagtca	actggtttcc	aagaatgggc	tccctctcat	caccttatga	54660

cagtaatcct	ctgtccaaca	gccaaagagg	tccgtgtggg	agggcttgca	gatgggagtg	54720
cgcagagccc	agctcaaagc	tccgtgactag	gctcctgttg	agtattcctt	tgattccctgc	54780
ttctgtcttt	ttaaatcaat	ggagacaggg	gaggggttatc	tccatcctcg	gctcaagatg	54840
aaatgcatcg	ttcctcgttt	ttctcattcc	ttcccaatgt	gtgtactgtt	aacttttagtt	54900
atgaagggaa	ttacagtgtc	ctgtgcatat	accaaggctg	tccaacctcc	acacctttgc	54960
tcaagctgtt	ccttctactt	gaaatgcctg	tttccttccc	ttctaattgc	atctttccat	55020
ccaggtagga	atcagctcct	tggttcatgg	agccttttct	gctctgtttt	actatgcatg	55080
gacttccttc	tgaattagca	gaggatgttt	cctagcttgg	tcttaaccct	tctccttttg	55140
tttgacctca	atcttactcat	cttaccat	agggtgtgaag	ctaattgaat	acaggatcta	55200
tgcttcactc	tgattttatc	tccacctgga	tagcatcatt	tttgaacac	aagcaggcat	55260
atgggagggg	agagaagttt	gggtccagaa	agaactggat	ttgaattcta	acctgtgtgt	55320
ttacgtgagt	acgttactta	accattaatt	acttcaatgt	atatttatta	agtacctact	55380
atgtgccggg	cactgtacta	agcaccaagg	atacaatggt	gagtaaagag	atgcagcctt	55440
caccatcacg	aaggaagaca	gatgttaatc	cattaaccac	gtaatctcac	aagaaaagta	55500
aaatgactaa	ctgataagga	caagcccctg	gagctacaag	aggggtgtata	cagggcatcg	55560
atccaataag	ggcagtgttg	cggggagatc	aggagccaca	cagagcctgg	gttgtctcac	55620
ttggaaaatg	gggtatcaac	cacctacttc	actaggtttt	taaaatcagg	ttaaatgagg	55680
taatactgtc	catgaacagt	atcttctgtg	ttgatgattg	attgaaacgg	agtctcactc	55740
tctcgcccaa	gctggagtgc	agtgggtgca	tctcagctca	ctgcaacctc	tacttccctg	55800
gttcaagtga	ttctcctgcc	tcagactccc	aagtagctgg	gattacaggc	agccaccctt	55860
atgcctgact	aatttttgtg	tttttagtag	agacaaggct	ttgcaatgtt	gaccaggctg	55920
gtctcaacct	cctgacctca	aaagatccac	ccacctcagc	ctcccaaagt	gctggggtca	55980
caggcatgag	ccactgcatc	cagccacttg	ccatgcatgg	catttaaaaa	tggtcagtaa	56040
atgttaccat	aatgaaggct	ggtaggttgg	ccaactgagt	ggtctgattc	agaaggaaag	56100
aagttagaca	tacgtgaaca	tttctgttac	ttgaagatcc	tcaggacagt	gactcctaga	56160
cccactcttc	atcacagtca	gctgggaagc	ttttaaaaa	atgcagacat	ctgaccttca	56220
cgctagacct	attagccaag	cagaagtttc	tgggcagggc	atctgcata	ttttaaaaa	56280
ctttaataag	gcagcctcaa	aattacagat	tcagcacgca	tttaccataa	ccactgaaga	56340
aatgcaaaag	tataaaaaa	agataaacia	caatctgtct	cctgctttct	tccctctcct	56400
cccctgcttc	tggaggcaac	aaggtcaact	atctgtgtgt	atctctttta	gcattccctc	56460
catcaatggt	cacataagga	tgctcacaga	taagcaccta	tgcggggggt	ttttttttcc	56520
ttgtaaaact	attcacatac	taaatacttt	cctcagtatc	ttgccttttt	tcaattcatg	56580
tcacagaac	atctcttcag	gtttatagat	acaggtccag	ctcttctttt	catagccata	56640
taacattctg	tagaatagag	aggacacatt	ttactcagt	tccgattgat	ggatatcaat	56700
attgttttca	tttctacaaa	tagtcaagga	ataacataac	tctgtaaaag	ttttattact	56760
tataggcgca	tttatgccta	aaggatagtc	tcaaaagagt	gaaactgata	aaatgtgcat	56820
ttttttat	taataggtat	ggacagat	gttctcaaaa	tggttgtggc	agttcaaaac	56880
accagtaaaa	caggggagat	atgtat	gaaaagcacc	caaggcgatt	ctgaagtgtg	56940
gcccaggata	agaaccattg	cccagagctg	tccagatgg	cccctgggtt	cctgaagtgg	57000
gtatcgggag	agaaatcttc	actgaatgaa	tgagtgggct	ccccagggaa	gtgatgaaat	57060
ggctccttatc	agccttgcta	tctccctctg	acagaggcaa	actctctctc	cctgggggaa	57120
gttctctcaa	ggcctctata	taagaagtct	ttgtgagagg	aagcaagaa	ggacctgggc	57180
tttgggaaga	tctaaagacc	caggaaggct	tctgggtggg	tgagtgtctt	ctctgctgtg	57240
gtggagctgg	tgacagttta	ttctcccagg	aggtccctgg	ctgtggctga	cagtttctgg	57300
agggctggca	ggcgtctacc	tgtggctttc	tttatatgag	gatgtcagca	ggggcagcct	57360
tcactctctg	ccttgacat	tcttctctgc	ggatgtgaaa	gtgctccttg	gctggggaaa	57420
ggagatggtg	gagacatgga	ggagggtgtg	gggtggtctt	tgaactctga	ggaggggaca	57480
taccttctaa	gtcctatgtg	ttcctaggaa	agccaataat	cattgtctct	cccgcctttt	57540
ttatgtcata	gactctgagg	gacccattaa	gtacaaacia	ataagcgtaa	tagtcccttc	57600
tttacttccg	ggcctgaagg	aaagccagcc	tcagccaccc	ctcagggttt	gctgcgttct	57660
gtttagaaaag	aggctccttg	gtcctggatc	ctggagcatc	aggagctggg	cctggcatga	57720
gcttttctgg	cccatcctga	tttctattca	ggccttcttt	ttctccacct	cactcccacg	57780
gtccccta	ggtgtgattg	tgatgtgtgt	gcatgtgtgt	ctgtgtgtgt	caatgacaaa	57840
ctgtgttctc	cgttgcagga	taaagccaag	atgaaactcc	ccttacttct	ggctcttcta	57900
tttggggcag	tttctgctct	tcatctaagt	aagtgttttt	tgccctcagt	ctttcttctt	57960
ctgttttttc	cctttctatg	gtagatgggg	tcagagttac	acacccaccc	ccttctttga	58020
tctgtctcta	tttctgaatt	tctgtgtgct	ttaaagggatg	gggactctat	ggccaggagt	58080
tgaaggatt	tctcaaggcg	tctgttatgt	ctgtggctct	ggttctactg	tgacattccc	58140
aattttgtcc	tttctccatt	atgcttactt	tgagcttact	gagtgccctc	tctcctttta	58200
ctctcttagc	atcgccatga	agtaggtggt	attgtatacc	catttcacag	aaatacagct	58260
gggtgagat	ggaaccagta	cccaagccca	tgactgcccg	actctaagtc	catgctctta	58320
accaccttga	cctgttcagg	cagcttgggt	tccctcata	gagactgggt	tccaggttcc	58380
ccttcccagg	cagagttgag	cactctgatg	cccagggcaa	ggtgtgagct	gtctgtgggt	58440
ctggggaggga	acaaggggag	atgtgaagga	aggacactta	gctatcctcc	ctgccagggt	58500

ctgagacttc	cacctttgag	acccctttgg	gtgctaagac	gctgcctgag	gatgaggaga	58560
caccagagca	ggagatggag	gagacccctt	gcagggagct	ggaggaagag	gaggagtggg	58620
gctctggaag	tgaagatgcc	tccaagaaag	atggggctgt	tgagtctatc	tcagtggccag	58680
atatggtgga	caaaaacctt	acgtgtcctg	aggaagagga	cacagtataa	gtgggtggga	58740
tccctgggtg	ccagacctgc	cgctacctcc	tggtgagaag	tcttcagacg	tttagtcaag	58800
cttgggtgag	tgccctatgg	ctgaggctga	ggtgggagca	tggaacgggt	gtgggatatg	58860
ccccagcat	tgctatcact	ggctcttttt	cccattgagg	gccctggggg	tgctcagtaga	58920
acctgagcct	cagagagggtg	ttggggtaag	aggggagggc	cacctacaaa	cagaagttgc	58980
atthtggctc	ccaaccttca	aatgggtgtg	gcaggggagg	gagggaatga	attgtgggga	59040
ctcaagaccc	atgtgaattc	atgtaggaag	gatgctccat	tctttgtcct	ttatcctgcc	59100
ctgtagttaa	cttgccggag	gtgctacagg	ggcaacctgg	tttccatcca	caacttcaat	59160
attaattatc	gaatccagtg	ttctgtcagc	gcgctcaacc	aggggtcaagt	ctggattgga	59220
ggcaggatca	caggctcggg	aagagaagtg	tgaacactaa	atgggggtga	cctgctgatc	59280
tcagccagca	ctcagcttgc	atcagatttg	tctgtttttc	tctgtataa	tctccagaag	59340
aaccagggat	agatggacac	ccacagacaa	cactgagggg	gctgcctggg	cattcagga	59400
agagctaaag	atthagaatc	aggaggtttg	ggtccaagt	cctttccatc	tctcactatc	59460
tagtaactt	aagttagctg	ggcatggtgg	tgcatgtctg	taatcctagc	tacttgggag	59520
gctgaggcag	gagagtcaact	ggaacctggg	agacagaggt	tgcggtgagc	cgagatggag	59580
ccattgcaact	ccagcctggg	caacaagagc	gaaactccgc	ctcaaaaata	aataaataaa	59640
taaaataaaat	aaaaaaaaaa	ttaaaacaag	accatgagtt	tgthtctcct	tctctaggat	59700
gagttggcaa	cccttgthtct	acctthtgtt	agggctggaa	ggacaagcct	gtcactggga	59760
tgcatagaat	ctgtatggtg	taattgccgt	ggatcagcat	ttcagatgac	taggacagtt	59820
cccatcatgg	tccagcaggg	aagggcccat	tgcccgggtg	gcagcagaaa	gagctggcag	59880
atacggggcc	aggtctgctt	ctctgccttc	cctctgcccc	atcccttctt	cccctcttgc	59940
tttctccagg	gtcgtgagc	acgttttcag	tggttgagc	gcagccgctg	gaactttgca	60000
tactgggctg	ctcaccagcc	ctgggtccgc	ggtggtcact	gcgtggccct	gtgtaccga	60060
ggtgaggtgg	ggctggggat	gaacgatgga	aaggtctggg	agatgggaag	tgccccaaag	60120
aggagatgct	acaaagagcc	tgacctttg	tgggagaggc	ttcctgggtc	ttttatatac	60180
tctgactcca	cagcagtggtg	tggtgggaa	aagaggccct	cctgtgggtt	gagttgggat	60240
ggacaagagg	ctgaaagtcc	ctttctgttc	tgccctcaca	ggaggccact	ggcgtcgagc	60300
ccactgcctc	agaagacttc	ctttcatctg	ttcctactga	gctggtccca	gccagcagtt	60360
cagagctgcc	ctctcctggg	cagctgcctc	ccctcctctg	cttgccatcc	ctccctccac	60420
ctccctgcaa	taaaatgggt	tttactgaaa	tggttttatt	ttctcctctg	atcgcgatc	60480
cactctgctt	agccctcatt	gaaacttctt	ccttatcatc	tctccccaca	ccacaacttt	60540
catagaagtg	tcagaagcta	ctactccttg	aggaggagga	tgagggtgg	agttgggtct	60600
atggagcctt	ttggagatgg	aggaatgggc	tcagctagtt	ctcttcatag	aacacctgat	60660
tactgggcac	ctgcatagtg	ctgccaggac	ctttcaaggt	tgtaggtaga	ctcccaatgg	60720
cccagtttgc	atctctgtaa	ccaaaggcct	ttctctctc	tctctccaac	cccagaactg	60780
tggttggttt	tatatgtaag	gaagttaaca	tgctccctgg	aacagtccac	aacattcagg	60840
aatgaatgta	taagtaccgc	aatccccggc	ccctcaagtg	gaataaatct	aacatgtatt	60900
gggcaccatt	tcccagtggc	ctgctgtggt	agttggcctt	attccatgca	tttttatggg	60960
ctgccttccc	ttcctcaact	gcattctctg	ctccttccta	ctctctgcaa	ctcccaaata	61020
aacacttgta	cgcaactccc	tctctcagga	tctcctctg	gggaaacctg	atataagaca	61080
gcttgccatg	cgtcagactc	tgaatgaggc	ctgggaatac	aagacatagt	cctctggcac	61140
ttgggatata	tggttatttg	taacataggg	acaaaaacat	ctactagttg	ttatcgctta	61200
ttgagcacc	acaacatacc	cctgtctgtg	gcaggcacct	tgccctagatg	acctcatgtg	61260
atcaataaatt	atgagcccta	ttttacagaa	ccaggctcag	agaagttagg	atctgtcaaa	61320
agacttgccc	aagactgaac	ctctaaatgc	aactcatatt	gaaattcaac	tctgtcccaa	61380
agcatgtttac	tttaaccctt	gtgcttttac	agctggctac	tctcccctta	tggtcacacg	61440
gggatgaagc	acggggggag	gaaagccaga	ctgtctcact	cttgggttca	tcttgggaca	61500
caggacacca	gcccagctgg	aggtgagggg	gctttaatca	gaggggaggg	agggaaggcat	61560
tctcaacccc	ttctgtacta	gggaggtcag	cagaagaaaa	taattcaatg	ttctaaagcc	61620
atthtthtct	ccagcattcc	tccaattcat	agatcttcat	atgggattag	gggctcagag	61680
aggggtgaaa	caagaactct	atthtthtgg	agtgtggtat	agagaaggga	tgtaacttct	61740
ctaaggtcac	atagtaagtt	gagaaagaga	gagaaatcaa	actcaggttc	atthtcaacta	61800
ttgttccaca	agaatctggt	gatttcaaa	atgggtggact	atgggttcat	ccctgtgggtg	61860
agtgtgtgta	ggatgcagct	gaggtggaac	tttcaactct	tgccctcttg	gactttatat	61920
tctggtgtgg	aaaggcattg	cttcccttat	ttcaatatata	acaacaaagg	gtaataatat	61980
ttcccattta	ttaagcattt	actaggtgtc	aggtactgtg	ctaaatgtta	ggtgaacttt	62040
gtcttggtcc	tcataaatct	ctgcccgtgt	gggtgtgtac	tttgacagaa	gtttgacttc	62100
cagttccacag	agatcttctt	tgggggagta	atatcaagaa	ggggcacgaa	ggaagctgca	62160
gggtccctag	tcccatcctg	tatctcgacc	taggcatgtt	tacattgggtg	cattcactgt	62220
gaagtttccc	tgagcagtc	actctatagt	gtgctttata	ggagcacatt	gtacatccat	62280
tgaaaaattt	ttcttgggccg	ggcacggtgg	ctcatgtctg	taatcccagc	actttgggag	62340

gccgagacag	gcggatcacc	tgaggctcggg	agtttgagac	ctgcctgacc	aacatggaga	62400
aaccccgctct	ctactaaaaa	tacaaaaaaa	ttagccgggt	gtggtggcac	atgcctgtaa	62460
tcccagctac	taggaggtt	gaggctggag	aatcgcttga	acctgggagg	cgaagggtgc	62520
agttagccga	gatcgtgcc	ttgactcca	gcctgggcaa	caagagcgaa	actccgtctc	62580
aaaagaaaga	aagagatttt	ttctttttct	taaaaagtaa	aaatcatgaa	ataaggggac	62640
tgggctaata	ttccaaaata	ttgggtttgtg	tgtgaatttt	cctctccagt	aagatactaa	62700
ctaagctctg	tgaaactgtt	tatctatggt	tctttatcat	tgaatccttg	gagttcctta	62760
cactgtgcag	agcacagagt	aggggctcaa	tcaacagtgc	actcattgct	ttttcataga	62820
caagggccac	cctcactcaa	ctcatgtgcc	aggcatagtt	ctgagagcct	tgcttaagct	62880
gatnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnggtcact	aatttggtat	63000
agagttaatgt	tcattgattt	cattttattt	tgtttctgat	ttttaaatg	tgttttactt	63060
gttttcttcc	tattattatt	ttattttatt	tgtaaaacat	ttacatatca	gacattttaca	63120
ttttcccaaa	ggtaaaactg	tgaaacaaga	tatattcaaa	gaagtttact	ttccctctct	63180
gtttcttgta	ccccttttcc	tcttctttag	gtaaccattt	ttattttttt	aaatataaac	63240
attgtgtagg	tgtatataca	tgtattagtc	tgttttcatg	ctgctgataa	agacctatct	63300
gagactggga	agaaaaagag	gtttaattgg	actttacagt	ccacatggct	ggcaaggcct	63360
cagaatcatg	gcaggaggtg	aaaggcactt	cttacacggt	ggtggcaaga	gaaaaatgag	63420
gaagaatcaa	aagtggaaac	ccctgataaa	cccatcagat	ctcgtgagat	ttattcacta	63480
tcacaagaat	agcgtgggaa	agactggccc	ccatgattca	gttacccctc	cccactgggt	63540
cccaccaca	atacgtggga	attctgggag	atataattca	acgtgagatt	tggtggggga	63600
cacagccaaa	ccatatcaat	acatttcctt	ctcttttaga	taaaaggtag	tatactgtat	63660
acactattct	gcagagtttt	tttttttttt	gatgtaactc	tatcctgagg	gtgctctgta	63720
gcagggacct	ctcatgcctt	ttaaccactg	cctgggtctc	cattacatgg	ctgcagcata	63780
gttgccacag	cattcctgta	ctgatgacta	tttggtattg	ttccagtctt	ttgctattac	63840
cagtagtggt	acaaagagga	tctggctaca	tgttcaggg	ggggaggggc	agatgtgtag	63900
cctgtcagga	gggtattgca	gtaatccatg	actgagttaa	tggtagttaa	aagctaggat	63960
gagtcagtgg	gggtggagag	aagtgggcac	atttgaatga	tatgtaggag	gtgaatgatc	64020
agcattattg	atgagtttga	gggtgggcat	gtggggaaag	gattcgagga	tgactcccag	64080
gtttctgttg	ggacagtggg	tggtatgtgg	ctcctcccct	ttttccaatc	ttccttgggc	64140
cttcgctgac	ttctgttggt	ttggcctaca	gagagcttct	ttttcctctc	tgttcgccca	64200
ggttcctcca	ctttggcgg	ggccctctgc	tcgacggtgc	cttcgctggc	cctgacatcc	64260
ctgctgtgcc	tggtgcttgc	cctctgtgcc	tcagtcccca	tcctccctct	ccagtacctc	64320
accttcctcc	tgcaagtgat	cagccgctcc	ttcctctatg	ggagcaacgc	ggccttccctc	64380
acccttgctg	aagtggcctt	ggggcgggct	ctgtggagac	ggacacactg	gggcaaagag	64440
aagctggagg	taaagaaatt	gggaggcaag	gcggggcctg	gaggcagtca	ggtgcgggag	64500
actgggtttg	ggggcaggtg	tggtgggggt	gagaccagag	gtggtgggaa	ggatagaaca	64560
ttcatgcact	tgagccttta	catctgcgg	ggcctctccc	tctgttttct	acctggtgaa	64620
ctcgtattca	tcctctgagg	cccactctg	tttctgttct	ccagggaaga	aatgggaatg	64680
tgtcttccct	tctttgtgcc	cttagtactc	tagtcttact	tcctttgcta	gtgctgtcat	64740
tgtctggcat	gccatccatt	tacatgctg	tcttttcttt	cctggtgcag	cctgcatgag	64800
ggtcctgtct	gtttttccag	ggccccgcat	gtgccttctt	ctgggttctg	tggttcaaat	64860
gtctgagcag	agctgaagag	ggaaaggcca	gacaggtgtg	gttggagggc	aggcctagga	64920
caggggagct	ggggacaagc	ggccgacagc	cccagaggc	caggcttctg	cttggaggga	64980
gggtccctga	agctcactgg	aaccctctg	gtttctctcc	ccagtttccc	ttcagagcac	65040
tttggaagc	tctttgggct	ggtgatggcc	ttgtcggtg	tggtgtctct	gctccagttc	65100
cccattctca	ccctcatcaa	aggctccctt	cagaatgacc	cattttacgt	gagtactggg	65160
aggatgggga	tccttgccag	gaggcctggg	ccttaggcct	tggtgcccc	aaatctggct	65220
gtgatggcct	gggtatgtag	catggtgcag	cttcccaaag	ggtctgtgtt	attcaagtat	65280
ttggggcaaa	agtatttgtg	tgtgtgggga	aacagacatt	ctggagttag	gtggggaatt	65340
ctcacgaaac	ttcaagcaaa	atcctgagac	ctcaaagggt	tttctgctt	gtggtgagt	65400
caggccacc	ctggcctctc	ccctaggccc	acacagggtt	tcacagttg	gcccaggga	65460
caggacctct	gtgctttcac	ctctgtgtcc	ttacacctgg	agggatgctc	tgaggctctg	65520
ctctaggagg	tggtcgtgag	tctcctgctc	tttgagaaa	ctgaggctca	aagaggttac	65580
ttacgtgttc	agaggcacc	gctaaggagc	aaaagtcaac	tttgaattct	gtgttttgac	65640
tactgcacag	ctctatttgc	ctcatttttt	atttttaaa	cagcaaatct	tagaatagga	65700
gtttaaatcc	atcacttgga	gaaaagaaag	actaaatggt	ttttgttttt	gttttgagga	65760
cacgatcttg	ctttgtcacc	caggctggag	tgagtggtgc	caatctcgcc	tcactgcagc	65820
ctcgatctcc	tggaactcaag	cgatcctctc	atctcagcct	cctgagttag	tgacactaca	65880
ggcatgtgcc	accatgcca	gcttatttta	ttttattttt	ttgatagaca	ctgggggttc	65940
gctatgttgc	ctgggctgg	tttgaattcc	tggtcctcaag	cgatccaccc	gtctctgcct	66000
tccaaaatgc	tgtgattaca	ggcgtgaacc	actgtgcatg	gccaaaagag	taaacttgaa	66060
atctgaggcg	aatgacttga	ttgtgacatc	aggtgacct	gtaatcagct	gtgtattcta	66120
gctggtgcct	ctaccagctt	cccattgtgac	cttgaacatg	tcattgaatg	ctcgctaggc	66180

ctctgtttct	ttatctgtga	aatgggcttg	atattcctcc	tctaccccaa	ccgatagtcg	66240
agaatgaaaa	gtaactgaaa	gtccttcctc	cagggcacca	tagtgtctgg	gtgaaaagta	66300
gaatataaac	tcggtagact	tctgggtccct	tcattggtca	tggaatggac	cagtgtctgc	66360
ttcattgagc	aacagttctg	ttgttcagaa	ttcctggatt	tcacctcact	tctgtctccc	66420
ctgcaggtga	atgtgatgtt	catgcttgcc	attcttctga	cattcttcca	cccccttctg	66480
gtatatcggg	aatgccgtac	ttggaaaagaa	agtcctctg	caattgcata	gttcagaagc	66540
cctcactttt	cagccccgag	gatgggtttt	ttcatcttcc	accacctttg	aggacctcgt	66600
gtcccaaaa	actttgccta	tcccagcaaa	acacacacac	acacacacac	acacacaaaa	66660
taaagacaca	caaggacgtc	tgcgagcaaa	gaaaagaatc	tcagttgcga	agcagattga	66720
tatcacacag	actcaaagca	aaggcatgtg	gaacttcttt	atttcaaaac	agaagtgtct	66780
ccttgcaact	agccttggca	gacccttgac	tccaggggag	atgacctggg	ggaggagtg	66840
tgtcaactat	ttcttttaggc	ctgtttggct	ccgaagccta	tatgtgctg	gatcctctgc	66900
cacgggttaa	attttcaggt	gaagagtgg	gtgtgcatgg	cctcagctat	gcttcctggc	66960
tctccctcaa	gagtgcagcc	ttggctagag	aactcacagc	tctgggaaaa	agaggagcag	67020
acagggttcc	ctgggcccgag	tctcagccca	gccactgatg	ctggatgacc	ttggcctgac	67080
cctggctctg	tctcagaatc	acttttccca	tctgtaaaa	tgagatgaat	tttgggtgtg	67140
aaagttcttc	ctggagcaga	tgtcctagaa	ggttttagga	atagtgcag	agtcaggcca	67200
ccccaaaggg	catgggagcc	agctgacctg	cttgaccgaa	ggatttctga	cagactatct	67260
ttggggatgt	tttcaagaag	ggatataagt	tatttacttt	gggcatttaa	aagaaaaatt	67320
ctctcgggaa	taattttata	gaaaaataaa	gcttctgtgt	ctaaggcaac	tactgtttcc	67380
atctctctag	gctttggggc	ggggctgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtatgt	67440
gtatgtttct	gaggaggccc	tacctgggca	tgagagggta	gggaatctgg	ctacacatct	67500
agtggtggag	ctggaccag	aggtggggca	ggaacctga	ctatgattca	ccccgctggg	67560
cctgggatgt	ggggccagag	acttctctcc	ccaggaaacc	ctctgcttcc	tcttctctct	67620
cacatcctta	actaacttta	gcagaacctc	actcctcact	acacaccccc	agctagaagc	67680
gctggatgga	atcagaaatt	cctagtttga	gtttcaattc	tgccccctcag	cagctgggca	67740
agccccctaa	ccactctgag	tcaactagtt	cccacctgca	aagtgcagtt	aatcatttct	67800
atctctgatg	gcgattgtga	gaatgtaaag	tcattgcaac	tgctagcac	atggtaggag	67860
cacatgaggg	tttgctcctg	tgtttactca	tgaccttg	ggaggacggg	ggcaaagagg	67920
gagaagttga	gggtgcagga	ggagagatgg	caggtgggtg	ggatgggaga	atctggggca	67980
cacctgctgt	ctcattccca	ccttgctagg	agagggacta	ggaaagaaca	gtgggaggca	68040
gggggatggg	ggtggaaggc	aggggggtgg	aggcaggttc	atccatccat	tcattcaaca	68100
aatgtttatt	gagcacctgc	cacgtgtcag	gccctgtcct	gggtgctggg	gctataaaga	68160
tgagaagg	tctgaaaccc	agctcttctc	tcttctgtg	gatgtcgggg	tgtaatttcc	68220
aggggccaag	agcctgggtc	tgaggggcga	caccaaagtt	ctagtgggtg	ctattagcag	68280
cgtttaaatc	taatggatgg	atttgggtct	gttacctgc	tcaaaagctt	tcagcagctc	68340
cccactgtcc	acaggacaaa	aatccagatg	ctagcctggc	attcaaggct	gtcactagt	68400
tgatctcaac	ctctccccct	ccctctttac	ctcctaccaa	cagcggggca	gagcccacc	68460
ctgtggacca	agattcccag	tctctgggtc	tgtgtgtgca	ccagttctc	tgctgggtg	68520
gctcacctg	cctcagcttg	tgaaatccat	ctgggtctgt	gggatcctgc	tcaaaatgtc	68580
atcttctcca	aaaatcatta	ctcaggcttt	ccagcatgtc	tgagtccctg	gcacttgggtc	68640
acacccttcc	tggtgactgg	catttgcttc	cacatcatga	ccctcccacc	ccttgccctg	68700
gcagcatact	ccaggaggca	aggtctgttc	tcgcctggct	ctaattaatc	tgtgtcttacc	68760
atccacatgg	taccagctaa	ttcttggtga	atgaatgatc	gttgaatgag	tggaattctg	68820
ttttggcctc	agaaccaatt	agaaggagcc	agaaaacac	atgggggtgg	gggaggtgca	68880
gtgtgtgca	gtggaaaaaa	acccttctg	aaatctcagc	tctgtcactt	actttgtcag	68940
ctctgtgact	ttggatggac	cacttctttg	tcagtatggt	gggagaaata	gacatgcctc	69000
tctgggctgt	tgtaaggatt	acaaattagg	tcgagtgcct	ggcatgtggt	gggttgaaac	69060
gatcacagct	agcattacag	atgatataat	aaagccaaaa	aaagatgcct	aatgtccacc	69120
agttggtgaa	cggacaaaag	aaatgtacca	tatttgggat	attatttggc	aatcaaaaaa	69180
agtactgaca	cctgctacaa	cacggatgaa	tcttgaaaac	attagactaa	gtgaaagaag	69240
ccagacacaa	gaaactgcta	atgattccat	ttaaatatga	aatatcgggc	cagggtgcag	69300
tggtctatgc	ctgtaatccc	agcacttttg	gatgccaaag	tgggcagatc	acttgaggcc	69360
aggagtctgt	gaccagcctg	gccaacatgg	cgaaaccccg	tctctactaa	aaattagccg	69420
agtgtagtgg	catgcacctg	taatcccagc	tacttgggtg	gctgaggcac	aagaattggt	69480
tgagcctggc	aggtggaggt	tgcatgagc	caagatcgtg	ccactgcact	ccagcctgga	69540
tgacacagtg	aggttccgtc	tcaaaaaaaa	aaaaaaaaaa	ggaaaaagaa	aaaaagaaat	69600
ttccagaata	ggccaatctg	tagaggcaga	aaagtagattc	atgattgggt	aggcctgggt	69660
gtggaggcca	tggttagtga	tggtctaattg	ggaaggggtt	tcttttgggg	tgatgaaaat	69720
gggtggactt	atgggtatgt	aattataacct	caataaaaact	gttattttaa	ggaagaaaag	69780
atgcctggat	tcgccaggaa	gtgtacagta	gacttctgtg	agaatcagaa	atgatttctg	69840
gggaagatgg	gcgagaggag	agtaagtggg	agaagtgacc	acgtgcgcaa	ctctcatcgt	69900
tctgccctga	gagccttctc	cctgcaactt	tatttattta	tttattttga	aacaggttct	69960
cactctgtta	ccctggctgg	agtgcagtgg	tgtgatctca	gctcactgca	gcctcgacct	70020

gccaggctca	agcaatcctc	ctgtttgagc	tcctgagtag	ctgggactac	aggcgcatgc	70080
caccacatct	ggctaattctt	ttattttttt	atttttttat	ttatagagat	tggggagtct	70140
cactctgttg	ctcaggctgg	tgtaaaacgc	ctggactcaa	gtgatcctcc	caccttggcc	70200
tcccaaagtg	ttgggattat	gggtgtgagc	cactgtacct	ggcacctcct	gcaacttctt	70260
cctcaagtgg	aaccaatgag	gaagcaagca	actcagagct	tcacaagtt	ttgatttcaa	70320
tcagcaacgg	gcttccaatg	caaccttctt	ctcctgtaac	cagcctcagt	agagaggaac	70380
tggaggtgaa	ttggccccc	tcacaccccc	acagtgccaa	gctggggcct	tccatcaggg	70440
ggagaacaca	tgccgtgtaa	gggacagcca	acagcataaa	ataggaattg	tgtgatgatc	70500
cctttttaagc	ctattcagcc	cagggaagtg	catatgatca	gccccatttc	atagatgaag	70560
aaagtcaagg	tcacccatta	gcacattgtg	gggctgggtat	ttaaaacagg	tctgtctggc	70620
tcccaaggtc	acattcattt	agacattacc	tttactttac	atttcttctt	cttttcttct	70680
tcttcttctt	cttcttcttc	ttcttcttct	tcttcttctt	cttcttcttc	ttcttcttct	70740
tcttcttctt	cttcttcttc	ttcttcttct	tcttcttctt	cctcttcttc	ctcttcttcc	70800
tcttcttctt	cttcttcttc	tcttcttctt	cttcttcttc	ttccttctct	tcttcttctt	70860
cttcttcttc	ttcttctctt	tcttcttctt	cttttttttt	tgagggtggg	ccttgaactc	70920
tcttcttctt	ttgcccagg	tgaatgcagc	atcatcatac	ctaaatgcag	aggcatgagc	70980
ctggccttaa	gcaatcccc	tgccctggcc	tccaaaagt	ccaagatttc	acagacgaga	71040
caccatgccc	agcctgcatt	tattctcttg	taagaaagat	atcatttaaa	gagataatca	71100
aaataaagag	ggacatgaaa	aagacgcac	accattaatt	ggaccactca	ataatttaaa	71160
tgggttaacat	gttggtatgt	tccctcccgt	catttgactg	gatgtatgtg	ttttaaatgg	71220
tgatctcata	agcttttctt	tatgtaatca	aatagtagcc	aaaaacatga	cacctccctac	71280
ctgctcacaa	ccccatctcg	tggttctggc	acgccttgtt	tatccccatc	gagggcgagc	71340
tccctttccc	cttccctggc	tgtgtggggg	tcctagatga	cgggtgagcca	ctgttgtcaa	71400
ttggtcagca	gattggagag	tgcaataaat	aaaaacactc	agaaggcgag	tacccgctaa	71460
gtgggcttat	cacaaaagag	caccttgga	tattccagag	aatgacctca	ggtcctgccc	71520
tcactatcca	taatctggtg	ctaactgtac	tttagctgaa	ggtgctggca	gggaaaatgg	71580
aggtgctgct	aagaacactt	ctattctgtg	agaatcagag	atgatttcta	tctgcctga	71640
gcgagaggga	gtaagcagga	gaaacaaccc	acaggcacag	ctctcatctt	cagatagggt	71700
gagccttccc	cctggccacg	ggttttgttt	gtttgtttgt	ttgtttgttt	aagcctcgac	71760
ctcactctgt	cacccaggct	ggagtgtagt	ggcaagatca	tggtctactg	ccagtgggta	71820
ctcccaggct	caagcagtc	tccccaaatt	caaagcttgg	agtgtatggc	aaatgcttca	71880
tgtctaggag	ccctttttcc	tgccagcccc	tcaggggatt	gatgactctc	gaaacatttc	71940
ggtgtgacat	gggcacagca	gtgagtcatt	cctctgacat	tctttgggaa	cagacaggac	72000
catccaggct	tcagggcata	agatccagtc	ctctgggtgat	aaggagttca	accctccagg	72060
aatgtctgag	tgtatcttaa	accaggagcc	atggcttgtg	ttcacaccag	acagagtctc	72120
gattttgagg	tgttttgttt	gtttgtttgt	ttgtttgttt	gttttttgag	cttcgcctcc	72180
tctctgtcgc	caggctggag	tgcaagtgga	cgatctcagc	tcaactgcaac	gtgtgcacca	72240
cggttcaagc	gatttctcct	tctcagcctc	ctgagtagct	gggactacag	tttgacagg	72300
ccacacccgg	ctaatttttg	tatttttaat	agagactgtg	tttcaccatg	actgggatta	72360
tggctctgat	ctcttgacct	cgtgatccct	ccgcctcggc	ctcccaaaat	cccaaaccatc	72420
caggcatgag	ccaccgtggc	ccgcccgaat	ttgagttttt	atgttctaat	aggcctcgag	72480
tgctcacagg	ccctcagca	tattctttcc	tggttccagt	gtcacctccc	tgatcagctg	72540
gctggctaga	gcagtagggg	gtgtgggaaa	gctctgggct	ttgcaggcac	ggtaccacgg	72600
tgtgacctta	accaccctga	acctcagttt	ccctcactgt	aatggaaaata	agcttgggct	72660
cagtttgggt	caaggactag	agagtaacct	tgggaaataa	aggtagcagc	ctggcactta	72720
ctggagatgg	actgtccaag	accaacttcc	agttcctccc	cacacaagct	atactttata	72780
gattcctggg	acctccgctg	cttcatctgt	aaaatggagt	aacaatagga	cctggcatgt	72840
gagttgtaag	gattgagtg	ctggatgaac	gtcaagcact	tcaaagggga	ggctgcaggt	72900
agtgagtgat	caatataaac	cacctggctt	gtagcagggt	tgctgtgtgt	cttgtggagt	72960
gttattagta	acatctgtgt	gcccttcaga	gcgtgcacca	cacttcacac	cgcttttagct	73020
ctggaatgcc	actattatag	ttcaggatag	aaaacctccc	tgcaagcact	aacttcaaag	73080
tgtctccacc	gaacaaaaca	acacaagttc	tttattactt	ggaatgggaa	aatgaaaaat	73140
gcaaaaaaaa	aaaaagactt	tcgagttacc	ccaaatctta	agccaaagtc	atggagagaa	73200
atcaatcttc	atattcaatt	tttgcgatac	ttttgtctcc	ccagcagtc	tcagagccct	73260
tccaagcaca	cagaaatgtc	aattaccagg	ggcagggtca	tgaattcctt	ttagattggg	73320
gggctgggga	agagtgcagg	cagacagatc	tggtcctgtt	tatcacgttc	aatgcctacc	73380
tgctccttga	ggagtcatga	agcatcttag	tgctttgttt	tgctacctat	tgccacatag	73440
tcagagagta	ataaggataa	gtaaggctct	acgtgaaaag	tgctcgcccc	acattacagg	73500
taggtccttc	attaatggca	gctactaatt	ttttttgttg	acgcaaaatc	gacagagtct	73560
tcaagtagcg	tacatgacag	tgaacagatt	agtcagtggt	gctcaccgca	acttctgctt	73620
cgctctgtca	ccagggtgg	agcctcccga	gtagctggga	ttacaggcat	gtgccaccac	73680
tcaagcaatt	ctcctgtctc	attttttagta	gagacggggt	ttcaccatat	tggccagact	73740
gccagctaatt	tttttttggt	gtgatctgcc	caactcagac	tcccaaagtg	ctgggattac	73800
ggtctcaaac	tcctgacctt					73860

tgccatgagc	caccgcacct	ggctgtgaaa	cagtttttatt	gtgttttctgt	ggaatgtgtc	73920
ctacccaacc	tatagctaac	tcctatatgtt	ccctcagttc	tcagctcaga	tatcccttcc	73980
tttctgtact	gttacctagt	actggttttc	atagcaccag	gtacctctct	ggcatagagc	74040
ttgtcacagt	tgcagtttaa	tgtaccatca	taggatttta	aaaatattca	gttgtgtcct	74100
ccattaggct	ttcatttggg	aactccacgc	aggcagcagc	tgtatatatt	gtattgccta	74160
ctgtatcctg	agaactttgt	accctactta	gcacagaatg	gagggtcagt	aaatactgga	74220
catgagagag	agagagagag	agagaggaga	gggagagaga	gagagagaga	ttcaacctac	74280
aatcccagct	ctgagcttct	agttccctga	tggtagggac	tgtgatgtgt	ctcacacggt	74340
aatgagcact	tatgcagaag	aggctcagaa	aatttctcct	catggccaac	ggaagactta	74400
gagttctttt	ccaagctcca	ccgtttgctg	gcatgcaaaa	tttggactat	cacttaagtt	74460
ttccaagcct	tgctttttct	atccctaaca	taggacaata	ttcagcattg	ttgtttgttt	74520
gttgggggca	ccatgtttca	ggcacttagt	agattattgt	accaccacat	ttcaattgggt	74580
cctcctcaag	ccctgcaaca	tctgtgaggt	ggatcatcct	aacaactcac	agatgagcaa	74640
caggagactg	gggggatgag	ggaactgcc	aggaggtcca	gcttatgggc	agcagagcca	74700
agaatggaac	cagggctctt	tattttttta	tttttttatt	tttatttttt	aaccagggtc	74760
ttttaaacat	cgaggaccac	attcctttgtg	ctttccaaat	catcacctgc	cccatgcaac	74820
ttacagggtg	agttacatta	aacaacgtat	gttaaatggct	ttgtgctagt	tattcaccac	74880
cacaggggaa	gtgagtcacg	gacaagagtg	cagccgctcc	attcggtacc	tggctctgac	74940
acttacctgg	aaaatgactt	aaccattccc	aggatcagct	gtttgtctgt	aatttaggta	75000
gtttaatggc	acttgtgtcc	tagagttggt	tagaagggtg	aataatatgg	agcacttaac	75060
atacttagca	cctagaaaca	cttcctaaat	attagtgtgt	gctgttgtta	tcgttattaa	75120
aatttctgct	taagatctca	tttcaggga	cccaactcaa	tctttgacaa	gcttaaacaa	75180
aaattgcttt	tcttcattta	ttcacttaca	cagcaaacat	gaattgagcc	tgtactgtgt	75240
ttccagaact	gtgcaggacc	agagaggcac	aggtgaagga	agcaaggctc	tggctctact	75300
ggggaaacag	caagaagatt	gctacaatga	ggtgggaaga	gggctggact	agagagaagc	75360
cctgattagt	gtccttgcta	cctttctctg	ggagagccaa	ggcaggcttc	ctggaagagg	75420
tgatccttgg	ctgaaacttc	gatgaagaaa	aggaagagac	gcagtgggta	gggaggaaag	75480
ggcattctgg	gcagatgaaa	tgacatgtga	caaaatatgg	gtgatcannn	nnnnnnnnnn	75540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	75600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnngca	ctccagcctg	ggctactgag	tgagagaccc	75660
tgtctcaaaa	aaattaaaaa	aaaaaagtcc	agaagaacat	ttgggtctca	ctctgtggcc	75720
caggctggag	tatagtggca	caatcatagc	tactgcaca	ttcaaactcc	tggcctcaag	75780
tgatcctcct	gccttagcct	tgaataaagc	ttggattaca	gatgagccac	cacaccagc	75840
cagaccatta	ttcataatag	ccaaaatgtg	aaaacaaccc	aaatctccat	caactgacaa	75900
atggataaat	agaatgggtg	ttgatccata	caatggagta	tttactcagc	aataaaaaaga	75960
agtcctgata	catgctacaa	ggatgaacct	cgaaaacatt	atgctaagt	aaagcagcca	76020
atcacaaaag	gctacatatt	acaagattcc	atttaaatga	aatgttcaga	ataggtaaat	76080
ctaactttta	tcacaggcaa	agctatgaca	ggaaatatag	gagtgggtgc	ctagtgtctg	76140
ggggcagagg	tgggggtgag	gcgagtgtgt	actgctaagt	gtacagagtt	acttttgggg	76200
ataaagaaac	tgttctgaaa	tggactctgg	tgatgggtgc	actactctga	acatactaaa	76260
actgtttaat	tatatacttg	aaatgggtga	cttgtgaggc	atggaaatta	tatcttaata	76320
aagctgtttt	acatattttt	catattttaa	aatgcagggt	gagggatgag	ccctctaaag	76380
agaagcagga	gtttgaggag	gttctaaata	ttgtgtgggt	ggtaactgag	catataaatt	76440
tgtcagacct	catcaaaatg	tatgatgtaa	tcttcaagaa	agttgatttt	aaagaaacac	76500
caccagcacc	aggtggagaa	ggcaggaaga	agttacacaa	ggggtaggcc	aagagtgggt	76560
gctcatgtct	ataatcccag	cactgcggga	ggccgagctg	ggtgggtgac	ttgaggtcag	76620
gtgttcgaga	ccagcctggc	caacatgggt	aaagcccgtc	tctactaaaa	atacaaaaat	76680
tagccagacg	tgctcgcgtg	aaccagggg	gagaagggtg	cagtgaagca	agatcatgcc	76740
aatgcactcc	agcctgggtg	acagagtgtg	actctgtctc	aaaaaaaaaa	aaaaaagtta	76800
cataggggac	agtggcaggt	gtcaagggca	ggcagggtct	ctcctatctc	caggataaac	76860
tcatagggga	cttagatgcc	atgtgggtcc	ctaatagcc	tccacttggt	tcttgcagcc	76920
actcttatgt	gtatcatttc	atgtcaggcc	tcttcttccc	aaccaccca	gccatcccag	76980
cctggctgcc	aacccacct	cctccagccc	ctgtcacccc	ataattgggg	ccaggaggca	77040
tgggagagtc	gccatctctc	ggtgccatct	gttgcatctt	tacagataac	catggctgga	77100
tccggcagat	cctgggggtg	agcagccgct	gttcagagca	gtgatcaaga	cctcccctac	77160
tgcccccctc	aaggaatcgg	ttttcttcca	tagccacatc	aggtgctgtg	caggaaggag	77220
ttgaaacgag	aagccaggag	caacgagaag	gacactaaca	tttattaagc	actgcagact	77280
ctcacagcac	tcccacggaa	tcgatattat	tatccccatt	ctgaagacca	ggcaactgaa	77340
gctcaatggt	taaggaaact	accgaagtca	ccaactgata	aaagtgtatg	aagctgggat	77400
tcaaatacaa	gctaaacttc	cttccaagct	tactccacaa	cacagaggtt	ggggaaaggg	77460
gataaaaaaga	gaggggagcc	caattccatt	tacccccagc	tcctgaggcg	gagcttgtca	77520
gcacagctct	ctccttccca	gaataggaag	atacccatca	gaggcaagtc	ctagacacca	77580
gcagtggtaa	ctccctgccc	caaggcagct	gcagacagcc	tatggctgta	gttactgctc	77640
ccaaagagtg	ttagaattcc	cactcccagc	ttcggggcca	ctcacacaag	gtgattgaag	77700

tggaaccag	agactctcca	caatgccctc	ctagagtaaa	tgaggctatg	taactttgtc	77760
caaagttagta	atttgaaaac	ctgggggctc	ccagctcctg	aaaagggaag	gatgtggggc	77820
cctttatatt	catactccac	tttgtgcagc	tctcccttgt	cttatgatag	ccctattaag	77880
aaattcctct	cccagcacgt	ctccttcaaa	gagctctaga	cctgaggctg	tcagaggctt	77940
aggactctgc	ctattagtcc	cagggtctgg	atgaccagca	ggacacctgg	cattcagtga	78000
ccactggatt	agataaatga	aacagtgggc	agagtgcac	ccaatctccc	cctgaagttt	78060
gaagaggctg	agaagtgagg	ctgtccaact	gctgacctg	ctttctgtcc	acctggccac	78120
ctaacttttt	ctggcttcca	cctgcccctt	tgccatccct	ccccccagcc	caccagcccc	78180
atttttcaggc	atacctgggc	acgtgctgga	atagaagccc	tcgttcttca	gaatgatcaa	78240
cagggagccc	cagcccagga	gtacagcaga	gaagaagagg	ttctccagca	cagccgtgca	78300
ggccatccac	cagcgcctcc	ggtacgcctg	ttgcagcgtg	ggggccatgc	tggccccgag	78360
cctgcacaga	aacagagcgc	tgggtgaagg	gccccccagt	ggccccaggg	aagggctctg	78420
catcatggtg	gcacccgaga	cctctcgggc	cagcccgcga	ggagcccctc	atggaggccc	78480
catagagccc	tgggcttccc	agccggtgcc	aaggagctgg	ctccgcgcgc	actagcagtg	78540
ccagaggctg	acgcggcacg	gggctcccgc	tgagccacta	tcggaaacaa	ggaaggtcct	78600
gtctgcgcgc	tgacagcttc	tagcaggctg	ccgggtttct	tcacccaggg	cagggcgctc	78660
agggccgggc	tgctggggag	aaagtccgca	tctgccagg	tccccagagg	acagcaaggg	78720
gcagagcgcg	ctctgaagca	ccgcggggcc	atgtccggac	tctcgcgcga	ggaaagacc	78780
ctagaagctg	gcaggaagaa	gggcaagttc	aaggctaccc	tacgacccca	tcttcagttt	78840
gcccctccaa	gacctctcct	tccctctggg	gccggggcag	agcaagccct	ccccctttcc	78900
gtatcagggt	acccacgacc	ctacagctct	tcgggccaa	ccaacagctg	ccacgtggag	78960
ggagacccag	gacgggctct	cctcggttcc	ctcgtccccc	gcgcgcccct	cactcactcc	79020
gcagggctcg	gggcaccagg	ctttgcacct	cggaacccgc	ttgccccctt	ccagccccgg	79080
gagggggctc	ggacttcggc	aggaagtctg	gcggctgctg	actttataag	ggcagcgggtg	79140
gcggatgggc	tggcgggcgg	gtgtgtttac	caaagggagg	gaaagagccc	cagctcccc	79200
cgccgcggcc	gctgcagcct	cggcgggagg	agagggaaac	cgggcagcgc	gggggcgggg	79260
agcgacaact	gggatgagac	cgaggaaagc	ggagaggaga	agggcaagaa	agaccagag	79320
agaggggagg	aagtaccagt	cacttcttcc	agggggactc	ggtattctca	tctgtgaaac	79380
ggggctttgg	gttcaagcgc	tccaggagg	ccgctggaac	tctggcaaac	gcgcagctct	79440
aagcagagga	agtgcagcga	gcggggaccc	gggaggaaga	gaagagtcgg	aggggtcaga	79500
gaaagaaaa	gggaaggacg	cgcttgccga	gatgggacac	tgtgccgcgg	gaccgcgggc	79560
gcaagtaacg	gtctttcctt	gggaagcctg	gcagtgtcgg	cgggagccgg	cctcgggtgc	79620
tctcagccga	cgcatagccg	gagaccctac	gcgcgcccc	tccccgccca	cgctgctcac	79680
ctccggctcac	cgccaatga	cggccagca	gctgcggacg	cctccgggag	cgcaacgctt	79740
tcgcggcgcg	tccggagtcc	cgtgggcccc	gccctgagcc	gcgcggcgcg	tggggtcttc	79800
tctgcgtgca	ggacccggcc	gccacggagc	tcagccctga	cagcccggtg	gcctcgcctc	79860
cgctgtctcc	tcggaagaag	cgggggaact	gggaacccgc	cgggcgccag	aggtctgcga	79920
agctgggctt	ggatgaagtg	gatctgcgga	ggtgatagtt	gtatttacac	gcgtccggag	79980
ctgcgccccg	aggtgggggc	gggggtctcc	ttcttttccc	ctccccctag	gtcgaagttc	80040
acgcgcacgt	gactgcgccg	ctggtcccgg	acactctccc	tctggcacag	ccccagcacc	80100
tacattttcca	ccctggaccc	ccatcttctc	ccccaaagcc	ccagactaac	atcaggcagc	80160
gccctctgta	tccttggtca	aaacaaagtg	cgattcggct	gaagccgact	gaccgcgatt	80220
cagggccgccc	ttgggtgggg	ttttgaactg	tgacgtgga	agcagtgttt	tccgagaggg	80280
agagtggcac	gggtttcttt	ggagttagtc	agatcgaggt	ctgagtcctg	actttttaac	80340
tgactaccct	gggttaccta	gggcaagtta	cctctctgag	cctcagcttc	ctcctcttta	80400
aatcgggtta	aaatggaacc	tacctaactg	cccaaaggaa	tcgcgattgt	gatgcaggta	80460
aaatgctaag	catagcattt	ggcatagtaa	gcataatgtt	aattgttgct	gctgtcatta	80520
tttcagaaga	cctggtgatc	ggatgcttcc	agatcaacaa	ttgattgact	ccaggtaaat	80580
ctctcagcct	ccctgagcct	cagtatcctc	atctgtaaaa	tagactacta	tgggtgagg	80640
taatgagaag	taatctcatt	acatgtgagt	ttaattgtgt	gttaagagtg	ctgctaagtc	80700
atgctgagct	taatacctag	gtgatgggtt	gataggtgca	ataaaaccac	atggcataca	80760
tttacctacg	taacaaacct	gcacattctg	cacatgtacc	ccagaactta	aaataaaaat	80820
aaaatttttt	taaaaaaaga	gtgatactgg	tggccagggt	tgggtggttca	tgcctgtaat	80880
cccagaactt	tgggaggcca	aggcaggagg	atcgcttgag	ctcaggaggt	cgagaccaac	80940
ctggacaaca	tggtgaaacc	ccgtctctac	aaaaaagaaa	aaaaaatagc	caggcatggg	81000
gggtgtgcacc	tgacgtctca	gctaccagag	aggctgaagt	gggaggatca	ctgagctgga	81060
gagatggagg	ctgcagtgag	ccaagatcat	gccactacac	tccagcctgg	gtgacagagt	81120
aagactctgt	ctcaaaaaca	aaacaagaat	gactacagaa	agctccaaga	aggcctcaga	81180
taaaagggaa	ccctgaaca	gatgagccac	caagccaaga	gaggaaactaa	tggctaccat	81240
agacagggca	ctttccaaaa	taaaaatact	gttattaatt	cctcaagaca	tcatggtccc	81300
atthaaacct	atagcttttt	cacagagggg	gcttgaagct	ggagctgagc	tgagcaaggt	81360
tagaggtagg	atgcagagtc	aggtcgccct	ggcatttaag	tacggctcct	tccattcctc	81420
ccagaaggag	aatggcaaga	gcaaaggctt	agctgtggga	atggcacaag	gagttctcgg	81480
tggccaaagc	acatgtcagg	ctctgatggt	ttaacttctt	aaaatgcaat	actgcctccc	81540

agaacttcca	gatcaaggtc	aaactcctca	gctctacaca	gggggaccta	gagtcaactt	81600
tctaagctag	gagagtcacg	gatccctttg	agaatacaaa	agacagtggg	cgcggtggca	81660
gtggctcatg	cctgtaatcc	caacattttg	ggaggctgag	gcaggaggat	cacttgagcc	81720
caggagttca	agacctgctt	ggtcaacata	gtgagacccc	tatttctaca	aaaaattcag	81780
ctgagcatgg	tggcatgtgc	ctgtagtctc	agttactggg	gaggctgaag	taggatgatc	81840
cctgagcctg	ggaggtccag	gaagctggag	tgagccgaca	tctcgccact	gcaactccagc	81900
ctgggtgaca	gagaccctgt	ctcaaaaaaa	aaaaaaaaaa	gaagaaatat	gttattgatc	81960
tactcttgac	aaaaatgctt	gtgtgaacat	ggacacacac	actcatcaac	attcacattt	82020
caaggttttc	atggaccctt	tccatgaggc	tctagtggtc	catggacccc	catggctgga	82080
acacttgctc	ttcctcatct	caaccacat	ttccatggag	ttggactgtc	tgtgcatga	82140
ggacacaggc	ctcatttggg	gtgttcattc	actgctgtgt	atcccagcac	ccagaacagc	82200
acctcaccta	aggggactc	agcacatgtg	cagtgaagag	tcagtcagct	ggtttcacac	82260
ctcccagctc	ttgcacctgc	tattccttct	tgtgggaatg	acagatttcc	ttcatttctt	82320
tttttttttt	ttttgacaga	ttccagctct	gttgcccag	ttggagtaca	gtggcacgat	82380
ctcagctcac	tgcaacctct	gcctcccagg	ttcaagcaat	tctcatgcct	cagcctccca	82440
agtagctggg	attacaggtg	cacaccacca	cctgtgagct	gatatttttt	tcttttcttt	82500
tcttttttcc	tgagacagag	tctcactctg	ttcccaggc	tgagtgagc	tggcgtgatc	82560
tcggctcact	gcaagctcca	cctcccgggt	tcaagtgatt	ctcctgcctc	agcctcccaa	82620
gtagctgaga	ctacaggcgc	gcaccacat	gcctggctaa	tttttgtatt	ttttagtaga	82680
ggcgggggtt	caccatattg	gacaggctgg	tctcgaactc	ctgacctcgt	gatccgcccc	82740
cgttggcctc	ccaaggtgct	gagattacag	gtgtgagcca	ctgcactcgg	ccattttttg	82800
tattttttta	gtagagatgg	ggtttcacca	tgttggccag	gctggctctg	aactcctggc	82860
ctcacgtgat	ccaccacctt	tgggcaccca	aagtggtggg	attacaggca	tgaaccactg	82920
cgctcagcct	ccttcttcat	ttctaagtga	ctcatccttc	acaactcagc	tcaagtttca	82980
cttctctctg	gaagctctac	tctaggctgg	attcaggggc	ttgtccacat	acccacccaa	83040
tactctgctt	acctctatgg	aagtccccac	actgatctag	aataatcagc	ttagttttct	83100
gcccccatcc	cgccccatga	gatgtacatc	ttgtgggggc	aggaaccacc	acgtggtagg	83160
tgatttgtgt	gcctgtgccc	tatcacaggg	cctggcgccct	aataagcttg	cggcccaacat	83220
ttgttgaaata	aatgaaaagg	gaatggtggg	aaaggaagct	gaaaaggtag	gctaaaatca	83280
gttttgaatt	acctctggga	ggccaaggac	tttcagtcctt	gcagggtagg	taacaggaaa	83340
ctcctggatt	ttgttttctt	ttggttttgt	ttgtttttaa	tgaagggtag	cgttatcgct	83400
aggtttttgt	gtttaattaa	tggagcataat	attggaaaagg	acagagacct	taaagcagtt	83460
aggagaccac	cataatagtt	cacattttgc	agccataaaa	aggaatgagg	ccaggctagg	83520
tggtctcact	ctgtaatctt	atcacttcgg	gaggttgagg	caggcggtac	acctgaggtc	83580
aggagtttga	gaccagcctc	accaacatgg	agaaacccca	tctctactaa	aaatacaaaa	83640
ttatccaggc	gtggtgtgtac	atgcctgtaa	tcccagctac	tcaggaggct	gaggcaggag	83700
aatagcttga	atctgggagg	cagagggttc	ggtgagccga	gatcgtgcc	ttgcattgca	83760
ggtacatgga	tgaagctgga	agccatcatc	ctcagcaaac	taacacagga	acagaaaacc	83820
aaacaccgca	tgttctcact	cataagtagg	agctgaacat	tgaaaaacaca	tggacacaga	83880
ggggaacatc	acacactagg	gcccgttggg	gagtgggggt	tggggggtaa	ggggagggaa	83940
cttagaggac	gggacaatag	gtgcagcaaa	ccaccatgac	acacgtatac	atatgtgaca	84000
aacctgcaca	ttctgcacat	ggatcctgtt	ttgttttaag	aagaaataaa	gaaaaaacca	84060
agaagaaaca	aacaaacaaa	aataattccc	attttaaaca	ataaaaaata	ggccaggcat	84120
ggtgactcag	gtctataatc	ccaacacttt	gggaggccaa	cgcgggcaga	tctcttgagc	84180
ccaggagttc	aaggccagcc	tgggcaacat	ggcaaaaacc	tgtctctaca	aaaaatataa	84240
aacaaacaaa	caaaatagcc	aggagtgggtg	gtgcatgcct	gtcatcccag	ctactcaggt	84300
ggctgaggtg	ggagaatcac	ttaagcctgg	gaggcggagg	tagcagtgag	ctgagatcgt	84360
gccactgcac	tccacctgga	gcaacagagc	aagattttgt	ctctaaataa	ataaataaaa	84420
taataaaaaa	cagagaagag	gaaagacacc	tgagatatat	ttccatatct	gaatcaatag	84480
gatttatcaa	cgttctcctc	tacccccaaa	actaattcct	tcctaaactc	tgttctcctg	84540
acactactca	taggttaagt	ataacagcat	tatcacattg	gctgtcatgt	gggctcctgg	84600
ctagaggctg	cttcacagct	taatggacaa	gagcactgag	acagggtggg	tctaaatcct	84660
ggctctgcag	ctgattattt	gtgtgatttt	gtccaaatca	ctccatctca	tgagcctcac	84720
tcttctagtc	tgttaagtgc	tgaaaaataa	agtatccaat	tcaattcatt	atttaatgaa	84780
ttatttagcc	taacaaatag	ctattataaa	tatttaggct	gggcacagtg	gctcacgcct	84840
gtaatcccag	cactttggga	ggccaagggtg	ggcagatcac	ctgagtcagg	agtttgagac	84900
cagcctgacc	aacatgggtga	aaccccgctc	ctactaaaaa	tacaaaaatt	agctgggtgt	84960
ggtggcatgt	gcctgtaatc	ccagctactc	aggaggctga	ggcaggagaa	cgcttgaaac	85020
caggagacag	aggctgcagt	gagccaagat	cgtgccactg	cactctagcc	tgagcaacag	85080
agcaagactc	tgtctcaaaa	aaaaaaaaaa	aatctctgca	tgaaagaatgt	acataaaatg	85140
gtgcagccat	ttcggaaaaac	agtttggcag	gtcctcaaat	agttaaacat	agagttacca	85200
ctatagccca	gcaattccac	tcctaaatat	actacaccca	agagaattga	gaatatttgt	85260
taacacaaaa	atgtgtatac	aagtatttat	agctgtatta	ttcattacag	ctaaaaagtg	85320
caaacatccc	agcagtccat	cagctgatga	acggagaaaa	aaaatgtggt	ataccataac	85380

aatgtcatat	tatttggcca	taaaaaggaa	gtactgatac	atgctacaac	atggatgaac	85440
cttgataatg	ttattctaag	tgaaaagaa	cagacacaaa	agaccacata	ttgtatgact	85500
gcatttatat	gaagtgccca	gaataggcaa	atccacagag	acagaaagta	gattagtggg	85560
tgccagagac	tgaggggagg	agataatggg	aaatgtggaa	tgactgctaa	tggtatgggg	85620
tttcttcttg	gggtaaatgaa	aatgtgtgtac	aattagataa	tggtgatcat	tgtaaaactt	85680
tgtgaatata	caacatgctg	aattttatac	tttattatat	tttatttttt	ttgagacaag	85740
gtctcgctct	gtcaccagg	ctggagtgc	gtggcacgat	ctcagctcac	tgcaatctct	85800
ctgcctccca	ggctcaagca	atcctcctgc	ctcagcctcc	tgagtacctg	acactacagc	85860
atgtgctacc	atgcctggac	aatttttgca	tttttagtag	agacaggggt	tcgctatggt	85920
gcccaggctg	attttgaact	cctggactca	agtgtccgc	ccacctcagc	ctcccaagt	85980
gctaggatta	cagggtgaag	ccaccactcc	cggcctaaat	tgtattcttt	aaaagactga	86040
attgtatggg	gtgcgaatta	tatctcaatt	taaaaaaac	aaaacaaaac	aaaaaaaaaa	86100
cctttgcgtg	tgtcaggcac	tagggattcg	atgctgaata	agacacagac	cctaccctca	86160
gagaacacag	agcccagcag	gagagagtca	cagatgaatc	aagtgttaca	tcactctatag	86220
gaagcgccat	ggaagaaaga	catgggtgcc	tgagaacata	cgcttagaga	agggaaattc	86280
atctagactg	gggctcagg	aggaatcttt	cagggtgatg	cttgtgctca	gagttttcca	86340
tgtcagaatc	agtgaatttt	atcaatcctc	cagaggagga	aacagcaa	gaaaaatctt	86400
acaacaggag	gatgcggaga	cattccgaga	gctgatcaag	ggctggtgtg	aacaaagcac	86460
ataggatgca	gagcctgtgg	tgtgaggttg	cagctggaaa	ggtaaaacac	taattacatt	86520
ggatcttctg	agacaataaa	gagtatgcaa	taatctcaaa	cgaccgaaac	tgaccttctt	86580
cctccctaac	ttgcttgctt	ccactgttgc	ccgtatcata	aaagcaccac	cctcttctac	86640
ccagtggctt	aagacacgaa	actcaagtca	tcccaggctt	tctccccacc	tcactctcca	86700
catccagcct	atcagcgagc	ttgtgggtct	taccacgtaa	agacttctca	tctccagcta	86760
ctaccatccc	ccaagcccag	atcaccatca	gctcaggcct	ggactcctgc	aacctttcta	86820
accgggtctt	cccaatccta	cccccgcaac	atgaccccaa	tagcccatca	gaatggacta	86880
atcgagatgt	agatttgatc	aggccacatc	ccttgaaagg	cttctgtga	ccctcgggga	86940
aatgcacaaa	ctcccaatga	tggccctga	gtcctgtgcc	atctgggtct	gacctctgcc	87000
ctctgtgtct	ttgccatggt	aacctccttc	acaccattta	atactccatg	ctctctccta	87060
cctcaagtto	ttcctgggct	ggaacattct	ctgcactagc	ctagccaact	aaccttttag	87120
atcttttggt	tgtttggttg	tttggttggt	tgtttttgag	acagtcttgc	tctgttgcca	87180
ggctggagtg	caatggtgca	atctatctcg	gctcactgca	acctctgcct	gccgggttca	87240
agcaattctt	ctgccttagc	gtcctgagta	gctgagacta	taggcaccta	ccatcacgcc	87300
cggctaattt	ttgtattttc	agtggagggtg	ggttttcacc	atgttgcca	ggctggtctc	87360
gaactcctgg	cctcaaatga	ccacctcct	cggcctccta	aagtgtctgg	attacaagca	87420
tgagccactg	tgcccaggca	acacttcaga	tcttaatgat	catttccttt	aagtgcctga	87480
cctcttgtag	taactagcct	gactccagca	atgaatcctt	ttgcaatgta	acctatataa	87540
catctgagtt	tccttttgat	aaaactcatc	atatatttgt	tcctctgaca	gttcagaggg	87600
caagggcctt	tgcccacctt	cctcaccact	atctctcac	cacttaacac	agaactcac	87660
accaccatg	cctcctgcct	gacaaattcc	taaccatcct	tcaaattctca	ctcacctatt	87720
accttctggg	aggcagtctt	ccctgagcac	caagacaatg	ggacacattc	ctttatacac	87780
cctgctgaac	atctcttttt	tgagggcg	gtagagatga	gtgtctcact	atgctgcccc	87840
ggctgacctc	aaactcctgg	cctcaagcga	tcctcctgcc	ttggcctccc	aaaatgctgg	87900
gattacaggc	atgagccact	gtacctgacc	gcaactgggt	tagnnnnnnnn	nnnnnnnnnn	87960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88080
aaaaaaaaaa	aaaaaaaaaa	aaattatttt	aaattagcca	gacatggtag	tgcatgctgt	88140
tagtccaagc	tacttgggag	gctgaaggga	gaggatcact	tgagcccagg	aggttgaggc	88200
tgcaagtgagc	cgtgatcgta	ccactgtact	ccagcttggg	caacagagt	agacctcatc	88260
cctaaaaata	aagaagaaaa	tatggcaatt	tgactgtaca	tctctaattg	gatatatcct	88320
aaggatgaga	aaggaataag	gaaggacaga	aaaaaggaaa	caaagaagta	gcaacagtat	88380
ttagcaattg	tattgttatc	aagtaacatc	aatatttgta	aaaccagtaa	ttatatttaa	88440
aatactatat	atgtgtatgt	acattttacat	atgcatatgt	taggaacca	gtttatcaga	88500
ggaagagaaa	gggctacaaa	tgtaaaatca	aggaataaaa	aatttgaata	aaaatatcag	88560
tattaagtat	ttatgatatt	tttcttataa	aaaaattata	tatatgttaa	ctctatccaa	88620
aaccaaaag	cagtgacaac	ccaggagcaa	taaaaaacct	cagcatccag	actgtagtct	88680
ctaccatttc	caattaaaga	aacccagggc	tagttgggaa	aaatgacaat	ttcatgtcta	88740
gggcaagaaa	cacacctagt	gaaatggacc	tgaacattta	attgtgttag	aaagtaagga	88800
aactctctag	aaataatgtg	atttcatcta	aaagacacag	attctgggct	ggtaaaagtt	88860
tcaatggcca	aaggtgagac	aatttgagca	tcaagaagaa	tcatgacaga	acagattaaa	88920
acatgtcaaa	tattttttaa	aatgaaatat	tataaaagaa	acaattagta	gccatccctg	88980
aaggtcacta	gggcaccaac	tcatatttca	aactggtaaa	taaagtgtga	agccaagcat	89040
ttatttctgg	gtaacaaaat	agtaaggaat	gtttttcttt	ctagaagaat	tctagtgtat	89100
aaaagtagaa	gatagaaata	gaaaatcatc	cttttggcca	ggggcagtg	ctgtaatccc	89160
agaactttgg	gaggccaagg	cagggtgatc	actggagatc	aggagtgtga	gaccagcctg	89220

gccaacatgg	tgaaccccca	tctctactaa	aaatacaaaa	attagccagg	tagtgggtgc	89280
ctgtaatgcc	agctactcgg	gatgctgagg	caggagaatc	gcttgaacct	gggaggcgga	89340
ggctgcagta	ggctgagatt	acgccactgc	actccagcct	aggcaacaca	gtgagactct	89400
gtctcaaaaa	aaaaaaaaga	aagaaaagaa	aatcatcctt	ttgcgatcct	aatgaaataa	89460
tgggcctagg	cattgatcat	taatggctcc	taaaatcact	aagtatatgg	ttgatgggaa	89520
actttatagt	ggatggatca	gactcgcaat	gtctaaacca	gttgatcaat	cttaacatcg	89580
taacaagaca	acagacacca	ggggctgctg	acaggagaac	agaggaaacc	catagctcta	89640
ccactgagtt	attcacggca	aaaaaaaaaa	aaaaaaaaatt	aaactgcggt	tcctccaagc	89700
ttctaatacct	gttgtttaca	ggaataatccc	aaggaaagga	atacttttaa	atgacacatt	89760
aaaacaacgc	caaatccaaa	atatggggaa	atgaccaggt	ttcttcaaca	aataaacaag	89820
aaaaggtagg	ggggaggact	gttctagatt	ttaaagctta	tagaagacac	agcaacccaa	89880
tacactgcat	ggaccaggca	tggatcctaa	ttggaacaaa	ccaactgtaa	aaggatgtat	89940
ttgaaatgat	tggaggaatt	tgaacagtga	ctgcacagta	gatgatatga	agaaattatt	90000
gttattttttt	aggtgtgatc	atgattttat	ggtgatgttt	aagtaaaaga	ggccttattt	90060
gttagagata	catgtacggg	tatagagaaa	tatttacgga	tgaatgata	cgatgtctga	90120
gatttgcttt	gaaaactcta	gcagggtggg	gagaagcagg	tgcatgggtg	ggggaaggga	90180
tagatgaaat	aagtatgcaa	aatgttagtc	tacttttgtc	cctcctgacc	cagcagggtta	90240
aaatacctca	gcatacctct	actcctccaa	ccagggtccaa	ggatcaggcc	aaaactccct	90300
gatgtggtaa	acagcctgac	ccctctctac	ctctctctct	ccagccactt	ccctaagatt	90360
ccccagtgct	ctgtgcccta	gccagcccga	ctcatctgcc	cagattcctc	aatgtttcac	90420
tctctcattc	accattttga	ccccactgtg	ctcctgggcc	actctccaag	gccccgctc	90480
ttcatctcct	ccctccttac	tcacctctca	ggtcttggct	taggtgccat	tcctctccagg	90540
aagccttccc	tgacaccaat	cccatcctca	cctagaacag	attatgtgcg	cttctttgtg	90600
ccccccatgg	ccccctgtgg	gtttgcttca	cggattataa	ctgcctgact	acctgccttt	90660
ctccaccctc	tagactgaga	acaccttgag	aaaaagaaca	catctatctt	gtctgtcatt	90720
gaatccctgg	tgtctggcac	catggctgac	acataactat	cactcagtga	ctagtgtttt	90780
aatgaatgaa	tgagtgaac	tagacagggt	taagaacaaa	agagaagacc	aggcatgggtg	90840
cttcacgcgt	gttatcccag	cattttggga	ggctgaggcg	ggcagatcac	ccgaggctcag	90900
gagttcaaga	ccagcctgac	atgggtgaaac	cccgtctcta	ctaaaaagac	aaaaattagt	90960
gggggatggg	ggcacacgcc	tgtaatccca	gctacttggg	aggctgaggc	aggagaatct	91020
cttgaaccca	ggaggtgcag	gctgcaatga	tctgagatca	caccactgca	ctccagcctg	91080
ggcaacagag	taagactcta	tcacaaaaaa	aaaaaaaaaa	gagcgagaga	gagcgagaga	91140
agatgtcatg	gggtaaatga	agacctccct	tcctgttccc	ctgaccagcc	cctgcctccc	91200
cccgcatctc	acctgtcttt	cttgccttct	tctggtactt	ctgtttaagc	cggtccatga	91260
gcaggccatt	ccagggggca	cacagcactc	cgaactgagt	gaaggcaaa	gcatttgtgt	91320
aggtgctgac	tgcagaagga	agagagaggt	tgggttgatga	gaagtttcca	aaactccctt	91380
ccaggcaggg	actctcccac	cttacccttg	tctgcatgtc	cctcctcccc	acaccatcag	91440
accctcctct	ggtgtgtaca	gcctgtctgt	gaggctctgt	gttcccagct	gggacatgca	91500
gatgggtac	ctcccagccc	taccacatac	ctcgtgccat	gtccccaccg	gccatgttgg	91560
tcagcaagga	gttgagagtg	ccaatgaaga	ggtagtgcca	caactgtatc	acagacagcc	91620
acaccagggtg	ccaggcaaa	cgccgagaga	aagcgtagct	ccagaaggag	cggagttcct	91680
gcttctgccc	tgcccctggg	gtctctaatt	gggagaggag	gatctgggcg	tgaattacga	91740
ggaaagtgga	caggtaggat	ggggagtgtg	gaggcttcaa	tggaaacattt	cagatccggg	91800
cccacttctc	acccttggct	cccagaactc	accggcgctg	ctctgcacct	gcctcggact	91860
cacacatccc	cacctccctg	ctttgtcatg	ctggccctac	caccttggat	gacctctgtt	91920
gttcttctct	attgaaatcc	gatctgtctc	tcacagcctg	gtcaatgaca	cttcttgca	91980
taataccttc	ctgatctttc	tcagcgagaa	aggtgaaagg	aacgacaagg	agaggagaaa	92040
gtcagaaggg	agaggagaat	gagtgtggat	actctgttct	aacctgctcc	tcagcacctc	92100
cctttctttt	gataccagta	tcctgagttt	ctttgggaaa	tcttctctta	ccctaattctt	92160
catgggtccag	atgggacctat	gaattcagtg	ttctgttccc	ctctcaagggt	taaccaatga	92220
gatggttcct	cctaacaagg	cagaccggcc	atgagtttag	attaggatgg	acttaatcta	92280
aaataggtcc	tacaccctgg	caagtccaat	gtcctccctt	gattttggaa	gcttcccaga	92340
accctattct	tctttcttta	aaataaataa	ataaatacat	gttttggatc	caattgtcag	92400
atggtaaaaa	taaaaaacaa	aaaatcaatt	ttattctgta	tatttaagat	atacaatatg	92460
aggtcatagg	atacatatag	ctactaagat	ggttactaca	gttaagcaaa	ttaacatatc	92520
catcatctca	catttctacc	tgttttgtga	caagagcagt	taaaaatctac	ttgttttagga	92580
aagtcccaaa	cacaatgcag	tttgatgacc	tacagtcttc	gtgctgtgca	ttagatctct	92640
aggcctgttc	atcctgtcca	tctgtcctct	tttgtccttc	gacctgcate	tcctatctcc	92700
tctcccaccc	cgttcttatt	tctactgtag	ctagctgcgg	tttgtgatgt	gtgtaaccaa	92760
agacgcagaa	cagagaggaa	ggaaggaag	cagtgtataga	gttgggacaa	taagagaggg	92820
cggaccagga	agacctggag	aaatgggggc	actgtaccag	acttagtgca	atggcatcac	92880
agaagagggc	agaaccgagg	agtgggggga	agggaaggca	acccatggca	ggcgggcttc	92940
aaggggtggg	gaagtgatag	gatgcgaaat	agagaaaaga	gggacagaaa	agagacgaaa	93000
gccttgagcc	ctccattaag	tgagagggtt	gggaagatgc	ctaaggccct	tttctgtgcc	93060

tgccctttcct	gattctgggt	ccctggggga	gctctggagg	tgaggggcca	ggaaaggcac	93120
aaggagaggg	ttgggtctgg	aggagagatg	ggttagccag	cagggctcac	cttccttcgc	93180
tgaaaggaac	tcctttgact	gtagctccct	gttttcatgc	tcagctgttt	ccttctcttc	93240
ctttgtgggt	ccattcccag	ggcacaggct	atggaaacaa	aagccccacc	agcaaggcca	93300
aggactgtga	gccgaacctg	agactcagac	tggagggaat	agcatggtga	atccccacatt	93360
ccaccgcact	ttggaatcac	cttttagcca	ctctgatgcc	caggttgag	accagaccag	93420
ttaaatcaga	atgtctggag	gtgagagcca	ggcttccttt	tctaagatct	ctatgtgaat	93480
ctagtgattc	taataagcag	caaagtttag	gaagcatgaa	aagagtaggg	caggccccagg	93540
ttcaaatecc	agctctgcct	cttcctagca	acagaaagat	ggctcagact	taacccttct	93600
gagcctcatt	ttttgcattt	agaaaatgga	gataaggata	tctcagagga	ttattgtgag	93660
gatgaaatca	gagagcacat	ggggtctgac	aattagtaag	tgagcagcaa	aggaatgccc	93720
ttcctctact	ccttgtggca	aatgactgca	aaaatgatca	catttcttca	cctcctctgt	93780
atttccccc	atttgaatga	gactgcagct	ctatttcccc	atgccctgaa	tctgggccag	93840
ccttgtgaac	tgcttcagcc	aaaagaatgc	agcagaagtg	gctgtgcca	ttccaagctt	93900
aaatctcaag	aacgcttgtg	catttctgca	ctctttcaga	accctgaaat	cacggtgtga	93960
atgagccac	gctggcttgc	tggaggatga	cagccacgtg	acccaggcat	ccctgtcact	94020
ccaaacctat	gtgagtggg	ccatcctagc	atagccagcc	cccatgtaat	cctccaaatg	94080
atcagatgta	tgaatgagcc	ctgtcaaaat	catctacatc	tggccctgat	cagcggaact	94140
agccagctac	ccacagactt	gtgaaaaata	ataaatgctt	aacatttttag	gctgctgagt	94200
tttgagatag	tttgttatgc	agcaatagct	aacagatgca	ctgctccagt	cctcctcctc	94260
tcctgtgata	ggtttgcctt	accctgtcca	tcccacccta	gggccaatga	ggggctctgg	94320
cccacaatca	ccagatagtc	cttaccata	gctgtagtgt	gggggcagtg	ggtatgggat	94380
gtgcccccg	ggcatcagga	ggaaagtgcg	tgctacatgc	caggtactgc	agacagagat	94440
gaagatgaag	gaggccctga	ggctgatgcc	tttttcataa	agaagctgca	gaaggagaag	94500
gaaaaagtca	gtgtcacacc	cacgttcata	gcagcactat	tcacaatagc	caaaggatgg	94560
aagcaaaacta	agggctccatc	agcagatgaa	cagctaaaca	taatgtgatc	tatacacaca	94620
atcgaatatt	attcagcctt	aaaaaaggaa	agaggcaacc	atgctggctc	acacctacaa	94680
ttccagcact	ttgggacgca	cgaggatcac	ctgagccag	ttcaagacca	gccttgacaa	94740
catagtgaga	ccctcacccc	ttctctagaa	aatttttatt	taattagctg	ggtgtggtgg	94800
catacacctg	tagtgccagc	tactcaggag	gctgagtggg	aggatttctt	gagcccagaa	94860
gtttgaggct	gcagtgaatc	atgactgggc	cactgcaccc	cagcctggac	aatgaaacat	94920
gaccttgctt	ccaaataaaa	aaaaaaaaagg	aaaggaaaga	aattctgaca	catgctgcaa	94980
catggtatgaa	ctttaagagc	actatgcagg	gccaagctca	gtggttcctg	cctgtaatte	95040
tagtgcttta	gaagaccaag	acaggaggat	tgcttgagtc	caggagcttg	agaccagcct	95100
gggaaacagc	aagacctcat	ctctactaaa	aataaataaa	taaatacagct	gggcgtgatg	95160
gtgcacgcct	gtaattccag	ctacttggga	ggctgaggtg	agaggatatg	attacatgat	95220
tacatgcctg	taatcccagt	actttgggag	gctgaggcaa	gcagatcacc	tgaggccagg	95280
agttccagac	cagcctggcc	aacatggtga	aacccgtctc	ctactgaaaa	tacaaaattt	95340
agtggggcat	gggtggcagc	acctgtaatc	ccagctactc	gggtggttaa	ggcaggagaa	95400
tcgcttgaac	ccgggaggcg	gaggttgag	tgagccaaga	tcctgccacc	gcactccagc	95460
ctgggcaaca	gagcgagact	ctgtctcaaa	aaaaaaaaaa	ggttaagata	gtaaaattta	95520
tgttatgtat	attttattgc	atacaaaaac	atcagcagaa	gaggcagggg	ctggaaacct	95580
gttttctaag	gagtcctagt	acaagccatc	acctaactatc	ctgtaagctg	attagggaca	95640
cctggttacac	acatgcccc	acccaccccc	agcacacccc	ggcagtagag	gagtcctcat	95700
acgacccatc	cccacagccg	gtggagccctc	ctcgtgtggc	tccccagaga	tcttctagcc	95760
cagtgccttt	tttcccccaa	cgacagcaaa	ggccttttgt	tcaaagaaaa	ttttacacaa	95820
aaattcatct	tacaaaacac	accaatgggg	agcttgccag	tcatctccct	ctttattctc	95880
cttggtgact	ggtatgacat	caaagagaat	ccctaagttc	ctcaacagct	cagtttgaaa	95940
accaccgacc	tagcccaacc	tcttcccatt	ttacagagag	tgacgttgag	gtccagagag	96000
gtgcagtga	ttgctcaata	aattgacaga	gttaagcagca	gcaaagtcag	attaaactaa	96060
gaattcctgt	tcctgctccc	tttcccttcc	caactctaga	gagacaggag	agaggctggg	96120
catggtggct	catgcctgta	atttcagcac	tttgggaggt	caaggaaggc	ggattacttg	96180
aggtcaagag	ttcaagacca	gcctggccaa	catggcgaaa	ccccatctct	actaaaaata	96240
caaaaaataag	ctgactgtgg	tggcacgcct	atagtcaccg	ctactcagga	ggctgaggca	96300
ggagaattgc	ttaaacccac	taggcagaga	ttgcagttag	ccaagatccc	accaatgcac	96360
tccagcctgg	gagacaaaag	gagactccaa	ctcaataata	aaaaaaaaaa	aaaaagagag	96420
aggaaagaaa	gatgaggcag	ccatctgggt	tctccagggg	aaggagggag	aaccagaaa	96480
gtgactctta	tgccaggagt	agaaaggctt	gagtgcctca	ggggctcagt	ctctgcataa	96540
ccctccaaac	ctccaaagct	tatgggacta	agctagactc	atgtctgggt	ggtgactgcc	96600
agagatcctc	ttctctgccc	ccataacctg	caggcagctg	caactgcctg	tgacctaaaca	96660
ctaagcccca	agagaagtc	caggttggat	ggcttgagat	ccacactctt	cccttctctt	96720
cactcagcca	tctgtggtgt	gctggcttta	gtcctccagc	ttgctgcctc	ataattgaag	96780
catggttgcc	acaactccag	ctatcacatc	ctcacaccac	aacattcaat	gaggaagact	96840
ttgtttttac	tctgctttca	ccttgcgtca	gggaagaaaa	gtccccctga	atcttccact	96900

atatacactc	cctttatctc	attaaaaagg	actggatcat	atgctgacct	ccacctatca	96960
ctagcaacgg	gtaaatggat	tgccatgggt	ggctttaatc	aatcaggatt	catccccctg	97020
gctaagcggg	tcaactgccc	gataaaactg	ttcgcaatga	ataagacaga	atggttggtg	97080
attgacctct	aatagccttg	gcaacagttc	atccccctgat	accccaacat	cagccactgg	97140
gacagctgga	caagcctctg	tgtctgcccc	tgtctgacct	actagccact	tgccaccttc	97200
ttgtccaaac	tagaagctca	cagcagcaaa	cgccccactc	taaaggtccc	ccagcctcta	97260
cccaacactg	gccaagcac	attatgacca	ctgccacaaa	agcttgggca	agtctgaaga	97320
aggggcttag	cggttacaag	ctcaggctct	agaaccgaca	agcctgggtt	caagttccag	97380
tatcatggct	actagctgca	gaaccttcaa	caagcttttt	aacctcagag	actcaaatgc	97440
ttcatctgta	aaatgggggt	aacacagtac	ctacctcacc	gagttgatgg	agacaaataa	97500
tgcaaggttca	caagacaagt	gtctggcata	tacaagtgcc	cagtgaatgt	aggctgttgc	97560
tataatttacc	ttataataaa	ggaagactgc	cgaggaagag	tcaaagtctc	cattgtacag	97620
agtgatgatg	gtcgaacggg	gttggccaaa	taggttccca	atctggggat	gataggacta	97680
gcctggatca	cttattttatt	catgaaacag	atacttcctg	agcaccagc	atgtggcaga	97740
ccctccttat	acccaaactc	accctccacc	gctagagctc	ccacctcagc	ttggggccaa	97800
cccatctgag	gcagccaatt	atagaaaagg	gtctctcctt	ccctccacct	tccggccacc	97860
ctgccgagtg	cctgggatta	gggaaggctc	ccacctgcag	gttgggtgatg	agaaacagga	97920
ttcccccaat	ggtgagcatt	ggcatggcca	ggaagagcag	cacggctgag	cctggaccat	97980
caaagtcaga	ggtagggttg	tgtcatagt	caaccctaac	atggagcccc	aaactctgct	98040
cccacctgct	ccaaattccc	aacaatcctg	gtatccaggc	cccaattcta	gccagcgctt	98100
cagcgtcctt	caaggggtttt	taggataccg	gccaaggcct	cccagatat	ctctgtggaa	98160
gtcttctgag	ccaccttctt	cccaaccaaa	gttggctcct	agtctgtggc	aggccaggaa	98220
gtctacagac	agaggcagag	ctctaagtga	agccacctct	ctcttccctc	agtaaaccac	98280
aagctgcctc	tcccttttcat	ccttgacact	cctggaaaag	aagacctg	actcaggtcc	98340
ctggctcaac	cctctagccc	attoccta	tcatgggtatt	ggccttgagc	ttcaatcatc	98400
tgtttaatgg	gaacaacagt	tccctgctct	cctgtctcag	gtgctatgag	aactgagtga	98460
gaaaaggacc	atgggtctttt	ctttgttcac	taaactctga	gcacttcttt	ggtggccaggc	98520
attgtgcttg	gcactggaaa	tgcaagatga	atcagatagt	ccttgccctt	aaatagactg	98580
acatgcaaac	aaatggttat	aacaggtctg	gtaagtgtga	gaccacagca	aaaaagctca	98640
agagctgggc	taggggaacc	cttgacaaat	tcttccctcc	caaaccagac	ttctgcccac	98700
cattattctg	gccacaacct	atgcctgtcc	tattatttgc	taaaatgttt	taagttgact	98760
cactttttatc	caaaaagtat	ctatttttaa	aggacacttt	atatcactac	tgtagatgaa	98820
aacactggca	ttacttgtca	tgaatagaaa	gtaactgtca	aaataaatac	aatgaaagga	98880
aaacaatgtt	attcaattgt	agctggatgc	atltgacctt	agaatgttca	aagcctaaga	98940
cctgctcttc	ccatcagttg	taaaatcaca	ctggccccac	atgaagacat	tctttcatga	99000
aatcagaagg	actgaaagag	aaataaaaag	ggaatagctg	ttctaccagg	tgattttgatg	99060
tttgtttagtg	tagttcacgt	agtatgcgtg	tgcccctaac	atcctcttaa	ctaccgtgct	99120
atacctttaag	aagcactgcc	aagagcta	tttagagtat	tcacacagtt	taccattcaa	99180
tttctgtctt	tataaaatgt	acatctctcc	tactactaaa	ggttgagagc	tcctttcaca	99240
atagagtcct	tatgggtca	atgctttttt	caaaactgaa	aagccctata	ttatggagga	99300
agaggaggat	tggtgctcag	acgatttgca	ggcacgagtc	aaacattacc	cagccaccac	99360
ctccacattc	agttgcttaa	aaatcattta	caggctttta	gagtagatga	tgctggtttg	99420
ataaggagag	tggtttgaaa	taattgggtt	gagggtgctg	gccatctcat	gagatctgtg	99480
tgaacaaaga	cactcagcct	ctgtgtttgc	ccagcatgag	tgcaacaat	ctcatgatgc	99540
tgctcagctt	agcatagctt	acacacacaa	gagtaatgta	ctttctttcc	taaaacaaaa	99600
attgagccac	gggtctaaca	ctaggaagga	atattgggag	gcactctgtg	gccaccataa	99660
ccaaggcaat	gacagaaaga	agagtggagg	atcaggaggc	ctgcacatca	ggccccacct	99720
ccacttgctt	tctctgtggc	catggacatg	tctttgcaag	gggtcctgct	gtggcttcag	99780
tttctcctct	gtgtaatggg	tggaagggtg	gtggaaaata	aaccagattg	gagttccaga	99840
cttaacagac	tggtgaaatt	ttaaaacaaa	gattttgagt	acaatagggt	tgtcaacttt	99900
taccttgcta	agtaaggata	tttgcaaaat	ggctcattcat	ataatcattt	cattaaaaag	99960
agaaagagaa	cattttaaca	cataggagaa	ggatgtaaag	gttttttgtt	gttgttttgt	100020
tttttagggg	tttttggttg	tttttggtta	gagtcctact	ttgtcaccca	ggctggagtg	100080
caatgggtgtg	atcttggtct	actgcaacct	ctgcctcctg	ggttcaggca	attctcctgc	100140
ctcagcctcc	tgagtggctg	ggattacagg	cgcgccaccac	catccagcta	atttttgtat	100200
tttttttagt	agagaaggga	tttccaccatg	ttggccaggc	tggtcttgaa	ctcctgacct	100260
cagatgatcc	acccgcctca	gcctcccaaa	gtgctgggat	tacaggagtg	agccactgca	100320
cctggccaga	tgtaaaagttt	tgaataaatt	ctactctctg	aagtaatccc	tctccatcat	100380
ccttgctttt	cacattttct	caataaactg	ttttcacaga	ccagcaatag	ctcaagatcc	100440
ttccaggatt	ctttcaagct	gcagatctct	gaataaacct	gtggtctgta	tatcttgctt	100500
atagccctct	gtcacacct	gccccagcca	ccagggtgct	ctgagcttgc	atccctccca	100560
cccacctgac	agcactcacc	tgcaagggtg	aaggctatga	tgagtgtggc	ggtggtgtag	100620
aaaaatctga	aacacataac	aggaaaagca	gaatattgtc	aaggaggagg	aaacctggga	100680
gaaaaaacat	gattctgctc	agccagccca	caagtgtagg	acttgaccgc	accctcagcc	100740

tgggatgcaa	cgggcactga	tgcctctgag	ccccaggctc	aaaaccaggc	gcaagaagcc	100800
gcgatgagat	tgagatgtgg	tccctgacctc	atggacagtg	catttttgctc	attctgaggg	100860
ccaaggctag	catggaaagt	cttggacaat	gagctcagct	gacgatgtga	ttggcttggg	100920
acttagccag	gacagaatgg	gcaaagcgaa	ggctctccca	cctggaagcc	ccaacagccc	100980
aaccccttgg	agaaaggggt	tagtgccctg	tctgcaaate	aaggccttga	gttctaactc	101040
ctcctcactc	tgtgaccttg	ggcaaggcgc	tgtccttctc	tgggcctcag	gagccttttc	101100
tataaaaaaga	aatgatcggg	ctgatctagc	tcagagtgct	atgatttcag	gactacagtc	101160
ccaaggttat	caggctccct	tagcatttgg	gggtcttgta	aggcatggag	taaaaaaaa	101220
aaagcaatat	cctaaggctg	gagaagaggg	aggggacaaa	ggaaggggag	gaaaggggag	101280
gtagcagggg	gccaaggacc	aagaaggact	gaggtacagt	cattctgcat	ccaaggcctt	101340
aaattgtaa	gactggcctt	tactctggct	gtttccggaa	aggcaggccc	agccagccct	101400
cccgctctctc	tctctgacag	ccaatctcac	atgtgcctcc	ctgggagcac	ctgctctgag	101460
ctgtatcagc	cccagcagg	ccgctgatta	ccactgagcc	tggccacaga	gcacgagatt	101520
aggatgcagc	aacacactgt	gtgtgagatc	acgtcccga	ccttctgact	catctgcaca	101580
ggaaaccccc	ccagtctccc	ctccagtcag	aaggagcctg	aaattccacc	agtggcaata	101640
ccaaagaaac	ttcctattag	ctaagccctt	agggagtgat	tggctgttgg	ggcggggagg	101700
gggggcggtg	aggaggatga	ggatgaagcc	tgggcaacct	ggatgtgagg	ctgtgcaggg	101760
gatgaggaca	aggatccttg	gggtgaagga	agagaagagc	aatttttaggt	tttgcctaatt	101820
ttgtaaccct	ggctccaagc	cagcccttac	aggaagtcac	cctggcctcc	ggctcaattc	101880
agcacgtgat	aggggaagcca	catttatgca	gagcagggaa	cgaggtgaag	aaatggaagt	101940
ggggctgtgg	tgaagtgggc	aagtctagag	agagtcccgc	tgcctggggc	tgctcctaac	102000
agctgctggg	agcgagctgc	aggtgtgggt	cctggcaggg	tggccgggct	gtctgactct	102060
ggatttctact	ccaagctagg	ctgctgcctg	aaggattcct	cttaccacc	tttgctggg	102120
ctggcctttg	ggacttacat	ggctatgagg	cgtgccacgg	tggctctgaa	ccggtcaag	102180
atgtagccag	tggggaatgt	catgaagttg	ttcatgaagg	accccagggt	gaagatgagt	102240
gagaacctct	catcctgggc	tttgagctct	ggagtagaaa	aaaggtctcc	catgcatccc	102300
agccttctct	ccaaatgagc	acacaggctg	ggctcccctc	cacctcagac	agcttgtcgg	102360
tcgcaaacctt	gtcccttaag	ctgagttgaa	atgtggctgc	ccctaattac	ccctcaggag	102420
ctggtgcctc	cctcccaggc	acttcccaga	tcaagtgggg	tgagagctgc	tgacccttcc	102480
tctcatcata	gaaagagggg	tgggcagggg	gcagagtcct	tctgtctcct	tgccaccacg	102540
tgggagccag	acttaacttc	cttagaaaa	tcataccctgc	ccttaccagc	ctgcctgtg	102600
gcattgcaa	tcggcccagc	atctgggtcca	cacagatccct	taaagtaatc	ttcattcttg	102660
aagacaaaca	ctagtgaagg	ccagccaaag	aggacgccag	caaagcccag	gcattccagc	102720
agcccagtc	gcagtgtggc	cacgtgcagg	ggcaggccct	ggcccggcat	gagcagaagt	102780
ggagtggatc	ttcaaattccc	actttgtcct	cctggacgga	tcacaggcgc	cgtaagcctg	102840
gcgtttgagc	acttggaana	ttcctctggc	aagccaagcc	cttccttttc	cgtagctctc	102900
tggttgtttc	aggcctgggc	aaaaaccatc	agcgggtgat	tctctggatc	ctgtagaata	102960
aagatagagg	ctgctggaag	aggaggcctg	cgggaaaggg	aaaggttagac	tagagtattt	103020
tgtgaggtgc	atgaagaggc	aggatgatca	tggccgctgg	cagcaaatgt	ggggaataaa	103080
tactccaata	catcatctta	ggcactgcat	ttgagtaacc	acgtggcaag	tagagaggca	103140
ggtcttgatg	gccacctgga	gtcacagggtg	agataaacga	cttacccaat	agctccccgg	103200
gagcaggtgg	agaagcggag	ctcctgcgct	cgaattctga	ataccgtccc	ctaaaaataa	103260
gacagcaact	caaccagggtg	cagaggcagg	gagatttcat	acagaagaca	caaactccc	103320
ctgccaagtt	ggtgttatct	tcagtttact	gacaaatgaa	caaaagctcc	catggatttc	103380
aggaacttgc	ccaaggtcac	agggtcagtt	tagtcacgac	gcaggccatt	ctactgccag	103440
aaataccccc	aactcccatg	accctcgctt	aggactcgca	aacctggtcc	ccgccgccc	103500
tctctgcctc	aacttctacc	aggaaagcct	ccggggggccg	ctccccgcca	gcctccgcac	103560
cccgtccag	cctgcggcct	gccctcccgc	cagaggagcc	cgagggggcca	ggccgcgctc	103620
ggcgccccat	ggcgcccgaa	aggggaccct	tcgccctacc	cgctgtctcc	gcgccggggc	103680
tctccgcgcc	ctttccgcac	gggccagggtt	cgcattcgcg	cctctcgag	cccctcccag	103740
tcccctgctc	gcctccgccc	cctcctgccc	gcccgggaagg	ggctggggca	gacctccac	103800
tctccatcac	ttccttcttc	ttttcccttg	ctcacagcct	cccgcgccc	ttttacctct	103860
ccctcttgaa	acttctccct	ctagaacccc	ctagaacccc	agcgggtgtct	ttccctccct	103920
cctcgctgcc	tttcagcctc	ccagccccct	tgccctctgct	tcccctaacc	aagttagttg	103980
aatgctgtta	ctcgctcagg	cccacttagg	gaaaatgtca	cacccagcac	ccagaggaca	104040
cacagacagc	acatgagggc	atagggacac	acacactcta	tttgtgcatt	ttgccttgac	104100
cgtggtgttg	gcagggaaca	tatttttctt	atttgtctac	cagcttaacc	gtctctccca	104160
gtttcacact	ccagagctg	ccaaaaaat	ccccaccaca	gaatcaggaa	gccaagaacc	104220
aggactgagg	gcttttcaga	aaccatcccc	tggaggactg	ccccatattt	tcactcccaa	104280
aaacccctta	gatgactccc	tgcctcacc	ccgcccccca	ggttctgaaa	gagccttccc	104340
gccagactgc	attgattaac	cattcattgc	cccttttttt	attaatcaaa	gacatatata	104400
attgctcatc	ggagcttggtg	atcagcgtga	ggccttacta	agcagctgcc	ttactatcct	104460
tccagcccag	agcacgtgag	ctgacgtctt	ctttggcctg	tgtggccgtt	tccttgccaa	104520
aagctcagtt	tggggagagc	ttcttgcgta	ttagatgcag	tctgcagact	cccaacccca	104580

gctacctgga	tcccctgagg	gccaggaac	tccagctatt	ccaagcccac	tcctcttttt	104640
tttaagagga	agaaatagag	gttacgata	gggacagcca	gaactgagga	ttttccagct	104700
caccaccaaa	gcacaaaaga	taaaagtctg	caaccaccct	agtgacttga	ctgaatggag	104760
gaaggggtggc	tggggctctg	tacccaagc	tactcactag	ttatacaacc	tgaggcaagc	104820
tctttggctg	ccccacctgt	aagacgagga	caatagtacc	ttaattatag	gaattgtcat	104880
aaaagaagta	taagatgggt	gtatgaggtc	cctgcatggc	gcaggtgcta	taggcagatt	104940
gtagggtagt	agattttcta	gtctgcagtt	atgtagacag	agccagagaa	gcagctctgg	105000
ggaggaattt	caaaggaact	tgcccacggt	cattctacaa	agctgcagta	ccttcccaac	105060
tctgaaacgt	atgctctcat	caccccgctc	taacaaacat	ttggacatta	gagaaaacaa	105120
gtcttttctt	aaaataacat	tatttatggg	agaaaatcca	caaaaatata	gcatccagg	105180
acaaacagg	cttaagatgc	aagattttct	attttactgc	aagacacaaa	gactctgaaa	105240
ttaatgcatg	ccctatcttc	tgctctggca	tacattttag	tctcctgggg	ggatcagtaa	105300
gtgtggaagt	agcaagggag	aaacagaaaa	aagtcaaagt	aaagagacag	attttagaat	105360
gttaatctgc	aggagcctgc	cagaaagatc	tagctcatgg	gctatctgta	catccaggac	105420
tgaagcacgg	gacacggggc	aggctcgcca	gggttctgtc	caccttatct	tgttacctct	105480
cttgactctt	agagcctcca	ctccacatct	ccatcaatg	tctgcagaag	acgtggcctc	105540
cactaacaca	agtcttactg	aactgatggg	acaggaaatt	agaatatcct	ctgaaccatt	105600
cccatgttct	ttggttcgaa	ttccagcagc	tagaaaaggc	agatgctatt	ctgatcactc	105660
tctgctggtg	ctccaatgag	gattaatgag	taacatcaga	gagagaagtg	attataataa	105720
ggtctgacgg	tgaccccgat	gtcttcatcc	ttttctcttc	gcctccttcc	tcacatctcc	105780
acaccttttt	ttttttaatt	gactgattgg	ttcaacaaat	acatgtggta	cctcaggctc	105840
tgtgccaagt	gccgggattc	gtagagaaga	gattcagtg	ctgctctcaa	ggggctcatt	105900
ctcttgtggg	agagacagac	aaagaaaccc	aagatttctg	gagtgtggga	atggctctcc	105960
aggcagatgc	tagcacagca	cattgaaagg	cacggaacct	caacaaaaca	ataacattta	106020
ggaaccagct	agagcacagg	gtggtgaaga	aagtggaaa	atttgaggcc	agcgtcgcca	106080
tctaagttag	ggcatthaaga	attcagccca	catcaatcaa	tcatgtccta	ttgatttcac	106140
cccttaatat	ctctcctatc	tatccgtggc	cactgctcta	tgcaagagtt	catcatctct	106200
cacagaggca	tcatctgctt	ccaagccatc	gccatttctc	tgcaagagtt	tatttccatg	106260
gttcccaactg	gatggcttca	cttaactgct	caaaacccct	ctgaggtcca	gtcaactggc	106320
tggttaaggac	cagtccaggg	tctggggatg	ccagccatga	gacattgctt	tgaggggaag	106380
agggaagcata	gaactggatc	tcctgcatcc	tactgcccac	gtaccaatgc	tggaggtggg	106440
tttcttcccc	atcatcagca	agtctggata	tccaggatcc	accctatgga	tgtttttatg	106500
gacagagtgg	gaagatggat	atgtttaggt	tagggaaaga	gggtttgcca	aagagggcag	106560
tataagttag	ctgcactcca	tcattcccct	ggcacaaa	atggctagta	tcctctagtc	106620
ctcaagagca	ccaccttcca	atgcagctcc	tgctgttcca	cagacctctc	tcctcaaact	106680
tcctctgaac	aacctcagcg	agggcaattg	ccactctctg	ggcagagtcc	agatattctc	106740
tcctaccctc	tgacatcact	ttctaaattt	gtatatgtag	ataaaactctg	agccattcac	106800
ataaagggct	ttgatttcgg	atacgccaaa	cacataaaca	aacaaaacata	agctttcctt	106860
tcacaaatggg	ctcatgtaaa	ttaaaatgtt	tggttttcca	cctacggtct	tgaaaggggt	106920
ttctacagcc	tgttttggaa	gtcagaaagc	aaaaggtaaa	tgcaaacatc	atttcacctg	106980
cagagaaaat	tctaactctc	ttgaggcagt	gccaaaaata	atacaagcac	actgctatcg	107040
agccaattac	tggtatctct	gagcttccgt	ctcctcatct	ataaaattgg	aatcgagctg	107100
tatggattaa	agataatgta	tgaaaactgc	ccaattagta	cactattaat	aaatagcagc	107160
tactgttggt	aacaaatatt	attgacttac	tggaatacaa	agaggaaata	aagtcacatt	107220
tagggagaat	ttcaaagtgt	tcctaacctc	aaaaagaaat	aaattagggg	gaaaaccact	107280
aagtaatggg	tgagctcagt	ttaccttgct	taagaagtcc	caccctagag	aactgatctc	107340
tagatgacac	ccaaatgcac	tcagtacaac	cccccaagac	tgtctgggct	taaggcaggg	107400
gcttggtattg	tcctgtaaag	tgtgggaagt	ctgttcatga	gccacagtag	acaggaaggg	107460
gatggagtct	tagagctggc	tcttcagggt	atctcctagt	gtgttcaaag	cagttctcag	107520
gaggggtggg	aactactaca	tagccaaagta	aatatgaggc	ctccttgctc	tggggagacc	107580
tttctcttta	acagaggtga	atctgaaagg	atacccaaag	aggcactgga	gggtgggggc	107640
cactctggcc	cctcagagca	gccagctcag	cttcagtggg	tgctagaggc	agcagaggat	107700
cagcctggat	cagcctccct	ttcacatg	agaaaacaga	gctccccac	cagaccagat	107760
actggaagca	ctggggcagg	cctaagagaa	agcagagccc	caagccccac	cacaccagg	107820
cttatgaggc	tactgtacc	caccctcca	agccccagct	ccacttctat	gttcatcaag	107880
caactgttta	ctggtaactg	cacttcccga	tagactttgc	tagaaaggaa	tgctcagtg	107940
cactgacaat	atctaaacct	gcaactctaa	ggactaggct	ggggaacact	gtgtcaacat	108000
ggaggcacgt	cctaccctcg	agaagaaaaa	taaggaaat	caataatacc	tgctgggcac	108060
tgagcactga	ctatgtatct	gattcgaagc	gctttttgct	taatctgtca	ctgaattcca	108120
cgataggtgt	tgttatttag	atctattatc	tgggccaact	gaggcctaga	gggacgaagc	108180
aactccccc	acatcaccag	gtagcagtg	caggactggg	atagaaacct	gatgctctga	108240
ctgaaactaa	tgcttttttt	tttttttttt	tttttttttt	tgagacagca	tctcactctg	108300
tcaccaaggc	tggagtgaag	tggtgtgatc	tcagttcact	gcagcctcca	cttcccagg	108360
tcaagtgtatt	ctcctgcctc	agcctccga	gtagctggga	ttacaggcgt	gcaccaccat	108420

gccagctaa	tttttttgc	ttttttgtg	agatgggggt	tcgccatgtt	ggccagactg	108480
gtctcaaaact	cctgggctca	agtgttctgc	ctgccttggc	cccacaaagt	gctaggatta	108540
caggcgtgag	ccaccatgcc	cagccagctg	atgctcttaa	tctgtgccct	accagcctt	108600
cctgggaggg	ttcccaagag	ctacacagag	catgagttct	ggaatcgggt	tgatgggggt	108660
accagtattt	actaatagga	atgaagatgg	gtaattcttt	cagacagcac	ccttgattaa	108720
aacaagagag	tagtgctgcc	tctctgtgat	tctgtgtctc	cctgccctgc	tcacacagac	108780
accacacca	cccacacgca	tgatcatgaa	aagaggaaat	ggatccagga	gaaggagacg	108840
actcctgagt	gaaaacaacg	gggtttttca	cattgagagc	tttgcccaac	accccaaaga	108900
tgaaaagagc	aggaaactgc	tggggccgat	tgaacactgg	acttttgttg	tggaanaag	108960
caaagggaag	ccggaagaga	ctggaacagt	ttccatgggt	ctggaggatg	gggaagtggg	109020
tagggattag	ctggaggag	aaggagaagc	tgggggtgga	ggggaacctc	cacttgccag	109080
gagagcacat	gtaggatggg	aacccagat	gatactcaag	gcatggcatt	agaccagaag	109140
caagtctgtg	gtgaaattag	ggaaggctcc	actgcgagct	gtagacagag	cactggacaa	109200
ggaagtggga	gacccagggt	ccagtcctgg	ctctggagcc	ccctgggtgg	gctgccccgg	109260
gcacctttct	ctcttggggc	ctccattcct	acctctgtga	agcgagtgtc	gaacctctct	109320
tagccctgac	ttgtgaaat	gctgggactc	tgtacagagg	ctgacattaa	gcagggtatc	109380
gtcgtgggtg	gctgcaatgt	tcctccagat	gctgcacggg	agagggcaga	aaaggcctat	109440
atggtgagtc	cgccctggga	gcctctgctt	ggaagctgaa	gtggcctgag	agtgaactcag	109500
aaaccacgga	agttcccggg	gctgatgggt	tcttatagat	tgtacatgca	gctctcctcg	109560
tgggctgcaa	aaccgcaaga	tgggctgtga	ccactctcaa	ggaaagagcc	ctatctgcaa	109620
aaagcattct	gccctccagg	tcttaaagca	aacacagact	caatccttat	tccttttaag	109680
acaaaattgc	ctcaggggca	tcagggagcg	agcaggcctc	aaatgtgtgc	ctttctagaa	109740
ttctcaatga	aagcacctct	ttgggtatta	ataatgacaa	cagtaatgac	agtcattttac	109800
tgagtgtctg	ttttgggaca	ggcattgggc	taagagctat	atgtaatata	tattattatt	109860
tgatgcccac	agccacccca	taaggaggcg	gaggtactat	cattatgcc	actttaaaga	109920
tgaagaaact	gaggcctcaa	gagatgaagt	aacttggcca	agtcactcag	ccagtaaatg	109980
ggaagagata	gacttcccag	tatccagagc	ccatgttttc	accattatgc	tgaagtacct	110040
cttttctgtg	gccaatgtga	tctgcctcca	ggaatcctgt	cttgatgttc	ccttccccat	110100
acagaagtcc	tctctgtgtc	ctcttcagcc	tgatagtata	tcttttcata	ccattctttg	110160
gacatctctg	ttatactact	ccaatggtgt	tcctccccct	acccctccct	gggagcttag	110220
ttgttgtgat	taagtatag	ggaaatgacc	cacactaaac	aaactcataa	gagactgatt	110280
gataaacctg	aaatgcaatt	tattaattaa	cactgagaaa	tgaaccacc	cagcagatgg	110340
gaatcctaag	gctgactggt	cagcaccaatc	tctttcagga	aggacaggct	tttgggaaag	110400
gaaatcaata	ccagaagggt	ctttgttgag	tacaaagtca	gagggaaggg	agttgatgga	110460
ttgacacata	ggtgaagctt	gacatacctc	tataaagcct	ccatcctgcc	aaggatcaga	110520
atatccaagg	cagggagcca	tctgggtgtc	ctctcctttg	gacagtgtctg	gatttttctg	110580
gatcctatga	agatcttacc	tttctggctg	catttatcat	gattgtggaa	ggctttttgt	110640
ttccttgttt	gcttagatta	atttctgcgt	atttaataga	actgaaaggg	aatttccct	110700
tgagagcttc	tgaagaggaa	taatcaatac	atactagtgt	tgttgccctt	tgcagagaat	110760
tcacttctgt	gttgtcactg	tatcctcatg	cttccctata	atggaggggac	agagatggta	110820
aaaacatgga	cttggaagcc	agaccgtctg	ggtttgaatc	ctggctctgt	tacttataag	110880
ctctgcaacc	tcgggcagat	tacctaagtc	agtttccccct	tctctgaatt	ggggatataa	110940
tagcacccac	ctcaacatct	gtcaagagga	ttcaatgagg	gaatacacat	aaagtgtcta	111000
gaacagtgtc	tgccatctgg	taagcagctn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	111060
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	111120
nnnnnnnnnn	taatttttgt	attttttgga	gagacagggt	tttgccatgt	tgcccaagct	111180
ggctctcaaac	tcccgacctc	aggtgatcca	ccgcctcgg	cctoccaaag	tgctgggatt	111240
acaggcatga	gccaccacgc	caggccccac	acacctttta	aacaaccaga	tttcattcat	111300
cgggaagtgc	ctgtggggct	ggtgtggaca	tgtgggtgaa	ggtggcactg	ggagaagtta	111360
ggattctcca	tgacctctgt	tactcatatt	cccacactcc	tcaaattagc	ctgagtctcg	111420
aggacagtct	gatggctggg	caaaccctgc	ggcaaaccat	tccccagccc	tgccctctca	111480
accagagtcc	ttccgataca	tgattctggg	cagctgttgt	taccctgtgc	ctccatgttc	111540
ttccagagat	atccatgcat	gcatcggcac	atgtgtataa	ttattatata	tattatttat	111600
cccacaagct	tttgacatca	atagtagcat	attattaaat	tgttctgtac	attattttat	111660
taactttgga	tctctggtac	tgctcaatat	aggaacatat	agacctggcc	aggcacagtg	111720
gtcacatctc	gtaatcccag	cattttggga	ggctgagatg	ggtggatcac	ttgaggtcag	111780
gagttcgaaa	ccagcctggc	catcatgggtg	aaacccccat	ttctactaaa	aatacaaaaa	111840
ttagccaggt	gtgggtggcag	gcgcctgtaa	tcccagctac	ttggggaggct	gaggcaggag	111900
aattgcttga	acctgggagg	cagatgtttgc	agtaagccga	gatcacgcc	ctgcactcca	111960
gggtaggcaa	caaagcgaga	ctctgtctca	ataaaaaaaa	aaaagggaatg	tatagacctt	112020
ctttattctt	tttgatggct	gtaggtggat	gttctaaaaa	ttgtgtaacc	aatctcctat	112080
tgataaatatt	taagttatgt	cttcagcatc	atatgaaact	tacaaacaag	gttgcatgga	112140
ctatccatct	gtaaatgtct	ttttgaacat	ttctagaata	attgcaggat	aaactcctaa	112200
aatgagaatt	tctgggtcaa	agaggatatg	cattttacat	ttaatagata	tttgtcaaat	112260

tgtcttccaa	agtggctgta	ccaattaaca	ccccgacctg	taatgaatga	gagtgccttt	112320
tttccccaca	ccctggagag	atgaaaaatt	tatgggcccc	ctttggagtg	catgggtggag	112380
gaagctgttg	gccgttatat	aaccctcgtc	attaataaagc	ctgggggtgg	gggggggagaa	112440
agagaggtta	gttagtggtg	gcaaacatac	aattagatag	aagtaataag	ttctaattgtt	112500
cgatagcaga	ggaggggtgac	tatagttaac	aacaatgtat	tgtatatattc	aaaatagcta	112560
gaatggagga	cttaaaatat	tccaacacat	agaataata	aatgcttgct	tgccggccatg	112620
ccaccctgaa	tgtgccagat	cttgtttgtt	cttgggaagct	aagcaggggtt	gaacctgggtt	112680
agtatttggga	tgggagaaat	gataaatgct	tgaggtgata	gatatacctaa	ataccctgtc	112740
gaacattata	cattctatat	atgtaacaaa	atatcacacg	tatcccataa	atatgtacaa	112800
atataatgta	tcagtaaaga	gagggctggg	cacgggtggct	cacatctgta	atcccagcaa	112860
tttggcaggc	caggggtggga	ggatagcttg	aggccaggag	ttcaagatca	gcctgggcaa	112920
catagcgaga	ctctgtctcc	acaaaaata	aaaataaaaa	cgaattagcc	aggcgtgggtg	112980
atgcatgctt	atagtcctag	ctacttgga	ggctgatgca	ggaggattgc	ttgagcccgag	113040
gagtttgagg	ctgcagttag	cctacgactg	caccactgca	ctctccagcc	taggcaacag	113100
aggaagacca	tgttttctaaa	agaaataaat	taaataaaat	aaataaaaaat	aaaaagactg	113160
aaaagcagag	tggtaaagaga	aaggactttg	gggctcaaca	gtactagcct	tgaacctgg	113220
ctgttactta	cccatcgtgt	gataagcaaa	tgccttaacc	cctgtgtgcc	tcactttctt	113280
aacatataaaa	atagaagtaa	aatcatacc	cacttcaagg	gtcattataa	aaagccaata	113340
gagataatgt	atataaagct	tctggaataa	tgcctggcac	acagtaggag	tttaataact	113400
ggaaattcat	tgtttgtagt	ggcagccttc	tgaatctgtg	tcctctttgt	ccactaatgg	113460
ctttgatctg	gatttggctc	aggcaagacc	tggggaaggg	cagagactga	gggcaactgg	113520
aggtataggg	tgttctgagc	ttccccagca	gagtgaggct	gggaaaggct	tgggagacag	113580
accaggcagg	tgctgataag	accggaatgg	gaggtggag	cataaggcag	ttcagttttt	113640
cccaaagggg	ggtgtacaaa	acgatctcgt	atgactcctt	tatactgtta	atgttttcat	113700
tttatcgcg	actgaaaaac	aaaaccaaca	tatttaatga	atgattccaa	ggggattcct	113760
gcttttataa	aaaatgctaa	agtaggcatt	cacatgttta	aaaattgagt	tgattttaat	113820
tttaaaatta	ctaagtcata	gtacataatg	tgtgagccac	agctatcccc	aaaatcatga	113880
tagcgataca	ttaatgactg	aagttcttta	aacatcaaca	tacaatgcca	attccagaat	113940
tcagctcaaa	ttctgcaatt	acacaggctg	gggttgaaac	ccagcttttt	tgctaactgt	114000
gtaaaattag	gcaggagggc	taacctcgct	gaatctcaga	tgtctagtct	gtaaatgaa	114060
gataatgttt	gtttttatct	cacagagttg	ttgtgaagat	tcaataaaat	cacaacatgt	114120
gaggatgatc	tggtctgtgac	acctgtcacc	ccactgatct	ccagagttga	ttcggtgat	114180
caggctggct	gggcagggtgt	cccctttctc	cctcaccact	ccgcatgcat	tcctcccga	114240
actgcacact	tggtcaaaga	ggaagacctt	tcctgtataga	ggaggaccat	tcttcagtca	114300
agggtatatg	agcacctggt	ctgtcctgcc	agaatctccg	aaggagctct	cagtaaaatc	114360
acaagatttt	attgtgcatg	gtagcatgag	cctgtaatcc	cagctactca	ggaggccgag	114420
ctaggaggac	tgcttgagcc	caggagtttg	agaccagcct	gcgcaacata	gtgagacctt	114480
gtctcaaaaa	aaagaaagaa	agaaagaaaa	gaataataat	agtaataaat	cacctgtgca	114540
acgtgtcac	ttctctcttt	ggaatgtagt	aagtgtacct	aataaatgtg	atcattgtta	114600
tcacacaggt	gagcacaggc	taaagcatct	tgactttatt	ctataagcaa	taaaagagga	114660
tttgttttta	cagaactcat	tatgttgtga	aaataatttt	ccaacattaa	caaagaacat	114720
tcttcaagta	aaaggaaaac	cacccatcat	tctcccaacc	ttcaataatt	ttcaattttg	114780
catattctcc	agactttgtc	aacatgaata	cttactttac	atggctcgcaa	tcagtgttca	114840
tgcaaattct	tttatcctga	cttttataaa	caaataatgat	gttataaacc	ggtctccatg	114900
tttctgcata	ttctttataa	ttatcatttt	gtgctgtcat	aattattgcat	tgactatggt	114960
aactgcagtt	ttcttaacca	tttactgtc	tggggaaatg	gaggataatg	ccagggtcat	115020
gcctggagct	ttttttgtc	tattgcatta	tattcttaag	atcaaatccc	agcagtgaga	115080
ttagtcatgc	aaaaagtaat	aataattttca	aggctcttgt	tataattttac	tagattgttt	115140
tccagagttt	tgcacactgc	tcccagagat	gtaggaacac	agacgtcatc	caaccttgcc	115200
agtgtctgggt	gatggtgttt	ataaaacttct	gctaattttaa	taagtatgaa	atgctatcct	115260
cacacggctt	tcatttctat	ctctttgatc	attaacagggt	tgaactatatt	tccaagtatt	115320
tgtttactct	ctgcataccc	tcttgggtga	agtagtcac	cacttccttc	acctgtttat	115380
ctgttgaaagt	cttgaggctt	gttttataaa	tgtgagcgag	cacttcagag	tcaatagaca	115440
ttaatgtctt	ccagccagat	ttggccactg	aggctcctga	gcaggggaa	gcatgatcaa	115500
aactacaccc	tggacagatt	aaattaattg	gagaaaatgg	gctgagaggc	agagatatgt	115560
gtcactggcc	tactgtgttt	gatcctatag	tgggggcctg	aactggggca	acggcctgag	115620
tccccacta	ccagttagcag	gaggctccat	tgtccccca	tattagagct	tgccggcactt	115680
ccatttgccc	cacctctac	aataccccac	atacatgtac	tcactctccc	ttgcaaactct	115740
agtggcttca	acccacagaa	tttaaggggga	aaggaaattgt	tctgtcctgt	tcacttactg	115800
cagaaatgag	aaaagcgttg	ttcacatggg	atcacctaata	gaagggatgc	catccccaac	115860
ggtgcctata	aggaaatggg	ggagggttgg	agagttgtgc	aaaatgcaac	agggaaatcat	115920
cagagtcctc	tgcccatga	tagagggttc	tcaaatataag	agagtctaca	gcaactaatc	115980
tcacggccac	tctaggcagg	gcttcccaat	gcttccccaa	ccccacctcc	atcctagact	116040
ttacccactc	tgctgaacac	agatgttacc	catagcacct	tgcacatga	ttgtttgatt	116100

agcacctccc	acagtagact	gtgtttctga	taggtcagca	acatttgctg	agcacctact	116160
ctgcagggt	gtgccagggt	cacaaaataa	acaaagccaa	agacaacatg	gaccctgaac	116220
tcagcaagtt	cagagtcaag	tgggataggg	aggctctctt	cactggaagg	taactccaag	116280
aaacaatggg	actcaacttt	ctaaccaaga	gaactccagg	gagctaaaat	tctgacttct	116340
ggttaagact	ggtgtggagc	ttcattaaag	aagaaaagat	tcacccagac	ttgagttcat	116400
agcctggctt	tgcagctttt	aagtcagtga	acctttgatg	aagttatgtg	acctctccac	116460
ccagctgccc	ctaacacctt	gcaggggcag	ggctggagtg	caaagggagg	caactgtacc	116520
acagcctggg	aggcaccacc	ccactagtgc	aagccgggca	acctctgccc	ccaaggcatc	116580
cctagcctcc	caactgcaag	catcaatctt	gcacttggaa	aggaacctca	cctttgaaat	116640
ctaggttcaa	atthagaatg	atccagctcc	ttgaagttct	atacagaaat	acagccagca	116700
gccaggcccc	gtggctcacg	cctgtaatcc	cagcactttg	ggaggctgag	gtgggtggat	116760
cacttgagga	cagaagtctg	agaccagcct	gaccaacatg	gtgaaacccc	gtttctacta	116820
aaaatacaaa	attagccagg	catggtggta	cacgcctgta	atcccagcta	cttgggaggc	116880
tgaggcagga	gaattgcttg	aaccaggag	gcagaggttg	aagtgaagca	agatcgtgcc	116940
attgactctc	agcctgggca	acaagagcga	aactccatct	caaaaaaaaa	caaaacacac	117000
gcagcttccc	tccacttccc	aaccacagct	ccatctcaga	caacaagggg	cctcatgtcc	117060
atgacatata	atatccaacc	aacatgtcta	aggccaacc	acaccctctc	caaacatctg	117120
ccccttggcc	accctttggc	catgggttca	tgcactggca	gaaaggtagt	tcagagaaga	117180
agccccaaaa	gggcccggaa	gtccacttgg	gctttttgag	attccagggt	ccaggataac	117240
ctaagtgtgg	tctagaagag	agatgcagct	tctgggaggc	acattccttg	gtcttaggga	117300
cttcttggcc	ccatggaggg	aaactggcta	gatgaggggc	aaagcagagc	cctctaaagc	117360
acagggctca	gggaaggact	ctttttgacc	agatctaaga	gcagcactac	ctctctgagc	117420
ctgtttctcc	atctgtaaga	aggggacatt	aatagactct	ccccgctaga	gttactctac	117480
atcagccagc	acacgtaagt	tcatgacatg	aagcaagggc	ttaatatata	ccggtgttac	117540
tataaataat	aggccaggcg	tgggtggctca	cacttghtaac	cccagcactt	tgggaggccg	117600
aggaggatgg	atcacaaggt	caggagtttg	agatcagcct	ggtgaaacct	catctctacc	117660
aaaaatacaa	aaattggccg	ggtgtgggtg	cgtgcacctg	tagtcccagc	tacttgggag	117720
gctgaggcag	gagaattgct	tgaaccgggg	aggcagagtt	tgcaagtggc	caagatctca	117780
ccattgcatt	ccagcctggg	tgacagagca	agactgcatc	tcaaaataaa	taaatacaca	117840
catacataca	tacatacata	catacataca	tacatacata	catacaatac	atggacaggg	117900
accctaaaaa	tgagacaggg	aaagagaaaa	acatgttctg	acaaccttgc	cctttatact	117960
aatttaggtt	ttcttgcttg	ttttagaaag	ggcctggaca	ggagccctgt	ttccctcagg	118020
ccaggcagaa	caaggtgtgg	aactcactgt	ggaagggttc	tgggtgacaa	gtgcagcccc	118080
gtccctccac	gtcccagcac	agtaggcagc	acgtgtctcc	attgactggc	tcaggagcag	118140
gcctgttgac	cagtggggaga	gctgaggagc	ccagggtggg	gtctgaagga	atccctagaa	118200
aatctgattt	ttccccaggg	cccacatcac	gtgccagag	ctgggaaagt	ggaggcagca	118260
tggtgatctag	ctgagaggct	ccatttttgg	tagcttctag	tttgggagtc	acagagacac	118320
ctggatgata	cgaagatgta	gctttgcagg	actctctaga	acatggagtc	caagatattc	118380
ccttcaatga	tgggacactg	aagcccacag	agggaaggtc	tgtcccagtt	actcagccat	118440
ctggaggcag	agaccaggct	agaactcagg	acttttaatt	tggaccagga	ttccttttac	118500
cacagtgggc	agccctagca	agtgccagg	aggggtggaac	tgtgaaggtc	atccgagggg	118560
tagtacacgt	gggtaggaag	tcatatctaa	gaactgacct	ccagacctgg	ctctgccact	118620
cactccttat	gagaccacag	gtgctgggtg	cagtgggtca	cacctgcaat	cccagcactt	118680
tgggaggcca	aggcaggcag	attgcttgat	tccaggagtt	cgagaccagc	ctgggaaaca	118740
tagtgagacc	cccacctcta	ccaaaattag	ccaggcgtgg	tgggtgtctgc	ctgtagtccc	118800
agctacttgg	gaggetgagg	tgggaggact	gcttgagcct	gggaggcgga	ggttgcggtg	118860
agccaggatc	atgccactgc	acaccagcct	ggatgacaga	gtgagacaga	atgacacact	118920
gtctcaaaat	aaataaataa	atgacagcag	atcatcattt	ttctttctgc	ctctagactg	118980
caatgcctat	ttctccaggt	agtcactagg	ataaaaagtaa	aaataatatt	atcagcattt	119040
accaaataca	gggtcagcta	ctctgttatg	ttctttcatg	ctttgtttct	tttaagcctc	119100
aaacaactct	atgagctggg	aacaagtatc	gtccttcttc	ctccatctct	atthattttat	119160
ttatctatgt	atthattctat	ctattcattt	atthattttat	tttgagacaa	ggtccttcta	119220
agtcaccagg	gatggcctca	aacttgagct	aggaactagt	gtcaccaccc	cccaatttct	119280
tttattgatt	gattgattga	ttgattgact	ggttaatttt	gaggcagggg	tctcgctaag	119340
tcgacagggc	tgggtcttgaa	ctcctagtct	ttaagcaatcc	gcccgcctca	gcctcccaaa	119400
ttgtctggat	taccaacacg	agccaccatg	ccagccctct	cccatcttct	gaataggaaa	119460
actggggttt	gaaaaggtaa	gcgacttgcc	caaggtcccc	tccttagcta	gagagcttca	119520
gagccagggc	acaaacccat	caaagcctgt	gctctcgccc	attgagccac	cggacctcgt	119580
acactaaccg	ccaagtgttc	tacacagtga	aggtgacaaa	gaggtgaagg	gaagagccag	119640
ggaggttctg	tggactcact	cggtgggtat	gcccagaggg	aagggggatc	ttgggtggca	119700
cattgagagt	agctgcgctg	ttagtaagtg	agaactcgga	agtccagact	catccagctc	119760
gtgccaaata	acccctcctt	ctacctggtc	ctcttcccaa	agccagctgt	tctccagaca	119820
atgggggtgg	cgggggcggg	tgctctcttc	cttctcaggg	aaaatccgac	gctgagccca	119880
tctccagaga	tcttggcttc	ccgtggggct	gcagatccac	ctagagccac	cagagggcgg	119940

gccagcactg	cggccaaaggc	ttgaagaccc	agcacaccaa	agccccggcca	agcctccagc	120000
ccagtgtcca	agagtccagc	cagaggccga	gtcctcgatc	tcaaaatgtc	taactgcaga	120060
agcccaactc	atgttcaggc	atgatgtgtc	tgattctact	gggacaatca	ttgccaccaa	120120
agaattactg	ccaaaatagt	aacgacatta	gctacctacc	acccctccac	acaccaacac	120180
acctcatttt	accaagcact	ttctcatgcc	tgggggtgcc	ggagacttaa	tgagcctcg	120240
cgatgaccgg	gtagcctcac	cgtagacatg	aggaaactga	ggcacaagga	aggggagtac	120300
attgtctagg	gtcacctgga	gaactctgat	ctccagactc	aaatttccaa	ttcgtccccc	120360
ctccccccaa	ccctaaccgg	agctaggtgg	ggtggggaca	gcaaattgtg	atggggggag	120420
gtaagagggg	tcagagtgc	ctacagagaa	gaccaaatgc	attgtggcac	ctactgtaaa	120480
atgagaccag	ccagccaccc	acccaccagc	cagccaccta	aaagtcttca	gtgggcacct	120540
gctggaaaca	cgacaatgga	tgacacgagt	tccctgccct	caaaaagctg	gtagtctagt	120600
tgggggtgga	gggggtgagt	cagcagataa	ttatgggaaa	ccgtgacacc	tgtataagg	120660
gcggggatga	gcagaggggc	tgcgactgcc	tggagccagg	gattcccggg	cggggcttcc	120720
ctttcctcgc	agctcgtccc	aggaggagga	gtcccccccc	agcttcgggg	ttccgcctgc	120780
cttggggggc	cggggtcccc	tcccacccct	ccccgaagag	cgcgggcccc	gggaaccgat	120840
gacagcacac	ctgagtcagc	ccgcccgcga	cccgcctctc	agcgtctgtc	tccgcattct	120900
gtgatatttc	gtcccccg	agccagcccc	actgcgctcc	ggaggcagct	cgggcaacaa	120960
acccagcgac	agattgtgcc	gcggctcatt	ccgggggaag	acgccaaacc	ccaccctgct	121020
accccccaaca	ctccctcccc	gccgcgcgct	ccaggccctc	ccccaggcg	caggccctag	121080
tcgggggtgg	tccctgggga	acgcagggtc	ctgtcctgcc	tcctggaaat	agggggagcc	121140
ctgggtgagg	gaagacggga	gccccagaga	cttttcttct	tgtttctacc	tgatccgaaa	121200
acgagagggg	cgggaaagga	aatTTtaggg	cacagagagg	agctgggggg	ccgagaagg	121260
ccgaaaatgg	aaccagcagg	gggcacccga	gagccgaggt	gccccggggc	cgggagcctg	121320
ggaatgaaac	tggggaagag	ggggagagaa	agggaggcag	agacaccgag	acacacagag	121380
acgagagaca	gagacgcagg	gagccccgcg	gggaggagga	gagagacgaa	gacacagaga	121440
gacagtgaga	aagacagaag	accgggcagg	gaaacagacg	agtagagaca	gaaaagggtcc	121500
gagagagagt	gagggagggg	gggaacagag	agacagagac	cacgaaatat	gagtaagagt	121560
cgggggagaa	aaccagagaa	atcgaatgag	aacgcgagaa	gaacgagaga	ccgtggagg	121620
agcagagaat	gaatgggaag	aataagacca	acatttatca	agagccgact	gtatgccagg	121680
caactgcattg	gaccctggca	cggataggaa	aggaggagcc	gcggcgcggg	cagcggggcg	121740
aggggcttct	gtgctcgcg	gagcggcagc	ccagggggct	cagcagcccc	ggcaccgccc	121800
cacctgcggc	tccagcagcc	ccaacccccg	cagcgtgcc	tggccaccgt	acccgaagcg	121860
gctcccccg	gggccccgag	cctatcctac	gccggggcg	ctccgcgagc	gcgccgggccc	121920
gagtcaagac	ctgggtcaac	cgccctgcag	cctttgtagg	gaagtgccta	ggtgatgggt	121980
ctgctgatac	cgctgtgac	caggccatga	agggccagag	gggctccagt	gagaccataa	122040
tccgccccct	tttaaaagg	ggtagaggaa	gttcacgcga	agccaacagt	cttctcccca	122100
gctttgggtc	ctctcctgca	cccccgcggg	agataaggtc	tccccctccc	gacacatcat	122160
acatacacia	aaaaacgcac	acactgcac	gcgcgcccc	ctcgcacccc	cttgtaaatg	122220
cactaagggg	catacacaca	ccgggcacat	atttctttcc	acccatcccc	aagatcgcaa	122280
gcgcaaaacc	tcgcacagcc	tcacgtttcc	caccagctca	gacatgcacg	ctggcggaact	122340
ttcagcggct	caccctgtgt	cacactcacg	tgcccccccc	cccgttcccc	caagcccgtg	122400
caaagggtaa	cgggcaagca	tccctgagtca	cacctgcaca	agcatccttg	cgcgcacgtg	122460
cacgctcata	tgactcagc	cttgacgcga	caaactcttg	catatactat	tcttatagtc	122520
gcacactggg	cttgaggtct	gggagtggaa	ggaaaagtgg	aatcttgagg	ctgtcccagg	122580
ggacagaaat	gctggaggct	gggacactgg	cgcgaggagc	gcggctgggg	cggggggagg	122640
gggtgaccca	gaagctcatc	ttctcctgga	aagttgggag	gggggaacag	gacaagtcca	122700
cggcgttctc	ctaaactacc	gcattcccc	aagaagggat	ttctctagaa	gagtggcgcc	122760
gcgaggacga	tcgaacacag	tccctccgggt	cgcttaagcg	gggggggagg	gggcgggggtg	122820
gaggggggtta	gaaagccgct	cccgcctcct	agtgttcgag	aaagggttaa	gtcggcaagc	122880
cagcaaacga	gggaggagcc	agcgagtgcg	ggaaggagtg	gggggtggtg	ggaagagctt	122940
cctcgctgtc	cccactctcc	ctcggttagc	agcctgggca	cacggacaga	cggactgacg	123000
gactctcgag	cggaacagcg	agctagcggg	gcgcggggcg	tgggcgtcga	cggccagccc	123060
cagccttccc	cgccccgtcg	cgccccgccc	cgtcccgtcg	gggcccagtg	ctcctcccga	123120
ggccccgagc	ccgggcggcg	cagggtagag	cgccgcggcc	cggccacgca	gccccgggag	123180
tccccggccc	tccccggagc	cccggggggtc	cccccggtgc	atccggcggg	ctcaggggagc	123240
gagtgaggagc	gcccctcccc	cgctgcccc	tcccccgagc	atcgagacaa	gatgctgccc	123300
gggtctcaggc	ggctgctgca	aggtaagaac	gccagcggcg	ggagagcgga	gggcatcctg	123360
gggagagaag	cagggcgctcc	cctctttcag	ggattgaggg	tggggcagtt	ggggagggtg	123420
ggtaacctgg	ggaaggggaa	aagctcagcg	ctggggccgc	gcccccgccg	ccagggtgtg	123480
tctcagcagg	agggcacttg	gctgggagcc	cgcggggcg	tgcgaggagc	tctgaccga	123540
ggtgggacgc	agggggcagg	tggacccggc	ccggagcggg	gagggaggct	caggttccgc	123600
tgtccccgct	ccacctgctc	cgggggacgc	tgaaggactg	ggccggctgg	ggaagcgccg	123660
actcagcaac	tccctcctgc	cggtgcctca	gcactttctg	gccacctggg	aagacaggag	123720
atgtgggtag	ggggctgtct	ggggaggtag	gagggcgaga	gggaaatcca	agtggccctc	123780

tctggttagga	gagatggagg	gcgctagaaa	gaggatagtt	ctactgattg	agtacacagat	123840
aaggggtgtgg	gccagagact	gggggtgggg	tggggagggg	tcagggggag	agggatagga	123900
aggagaactc	aaagatggag	aaagtgggtga	gggaagctca	aaggaggagg	gagatggagc	123960
gggggagggg	gagaaggaat	aaaggttaga	tgggaaaagc	gtggagggaa	gtgggaccca	124020
ggtgaagacc	aaggaagagg	gaaggagagg	aaagaccaga	tcaggggagg	gatgggaaga	124080
agactatgga	cagggaccca	gaatcctggg	atggaggtag	cgggaaagag	aatcaggact	124140
gggaccttgg	ggactggaat	ggaaaaggag	aatggaaaga	tcagaaacca	gagaaggatg	124200
gggatggtga	ctagagaagg	ggtatcagga	accggcgaag	agggttggag	acagggaaacc	124260
atggatggga	gaggggctgg	agaggaggga	agaggaggag	gaagagaaag	gctgagagag	124320
agggactggg	gattgggggt	gctgcccagg	gatgagacaa	agaggcttct	ggttaaccact	124380
tccacgtggg	aagccctcca	ttcccaaagc	gcctgcctgc	cacatttctt	ctctcagggg	124440
gtggctgggt	ggccagatgg	ggggtgcttt	gagctcaggg	ccctgggggt	ggctgtgagg	124500
gacagagggg	gaggactttg	gaaggggagt	gacagcctcc	gaggggtggg	aaacaggctg	124560
gctcctgtgc	tgccatttat	ttatccggcc	cggacgttgg	attctgcagc	cgctgccgcc	124620
accacggttg	ctgcttattt	tggggtgtta	cattctggca	gagtgagaag	ctgtttgcag	124680
cagctcctaaa	cctccgtcac	ccgcgtcagt	gcctccccag	gcccctgcgt	cactggcatc	124740
accaccacct	cctcccact	cctcagctcc	cacctcctca	gcccctgccc	cctcagcatc	124800
tgcccgcagg	cccagccct	tcctgaagc	agcccgttgg	gtgtggagcc	cttgcttctc	124860
gtctgggacc	ctgtgcccct	ccttcagag	cgagaggcct	ctgtgcctt	tccagggagc	124920
atcctttcct	gggaccactc	tgaccagcg	actctgccct	gtgggtgggt	agcctggatc	124980
ctgcccccta	ctttgggtcc	agttttcttc	tcctcaagtt	ccttcttcta	caggggcctc	125040
cgccccaaag	agtggcctgt	gggctgagaa	ctttgtttct	gagccttggg	actccaaggt	125100
ttgatagcca	gagtccctga	cagtgggtccc	tcagtgaaca	gatacttttg	gctctggaca	125160
cttcagcctt	ccgggatcaa	taccatgttc	tggcctctct	tggctccctc	ccctggtcag	125220
ttctggccat	atattctgga	caggggtcat	ctcttcttga	ctcccacatg	taatcactac	125280
tctagaacaa	ccgcaactgg	aagcctagga	ggtgaaagtt	gcagagagag	ctggagtccc	125340
ttccttgctc	tgaccctgaa	tagccaaaca	gactcagcat	tgtggctggc	ccagccctag	125400
gcacctgggt	gcaatttctc	tcctgtcttt	acctcaaggg	cagtgtctca	cacattcagg	125460
cgtgggtttct	ccggaggttg	tggccacctc	ttaaagaaaag	atcagagtgt	ctctctgaca	125520
tgggcttgat	gtccctcttt	tccaatctgg	gttccacctt	gtactagctg	catgacctga	125580
ggccactgtg	tcattgtttct	ggggctccct	tccttcatct	gcaaattggg	ggccacaata	125640
ttgacctcca	ggggattatg	tgtgttgtgt	tcaatgtata	aagaagttaa	cctgtacaaa	125700
tgcagtgcct	aggacaaaat	aggtgcttct	tggtttccct	ctaccctgct	gtactctccc	125760
ctgcagctct	agccatcccc	tgctgacttt	agaggagggg	gtgagcagag	aggggtgggg	125820
aggctgctac	aaagggtctt	cctctgtcca	tgaagttagt	gagggatgaa	atgaaggctt	125880
ctgagaaaga	caatgaaggc	gagctgtaga	gacctggtca	ggaggcctgg	ggtgctcaga	125940
aactcacact	tccccctccc	agccctcaat	ggtgttaact	atgatgtgag	gggtcggtct	126000
taggttgcca	ccgaggtatc	cccccttcca	gctctgatac	tctgtgcac	ttgccccagt	126060
ctccaccggg	aattcacaaa	atgaaggcca	ggagtggagc	cgtggctctc	gggagagaca	126120
ggaggcctgg	gcctggaggg	aaggagtggg	ggtgctgagg	aggagtgaga	acaggggggtg	126180
gggaaggggac	gtggcaagaa	agaaaagggc	acacactggg	cagggcaggg	actgagggcg	126240
ggggagagag	ggaaaggcac	agctctctag	tccccaacc	ccccagtccc	accacctctg	126300
ccctggagtg	ctcgctccag	ccccagcagg	cctggggcag	tgaagcccag	agccccctcc	126360
cctccccctcc	tccttgccctc	cagtgaagag	cgctgcgtga	attatggatg	agctccttgg	126420
gttacagctg	ctttgcacgg	cagtggcaag	ggcagagaaat	ggcaacagag	tcactgttat	126480
gcagcagctg	ttatggagga	gccccagca	ccgggtcgct	cttcagagag	cctgcaggga	126540
ccactatcat	gggctggggg	aggtgagccc	tggttggggg	agacatggga	acaagatgga	126600
aggagagtgg	ggaaagagaa	gagaagtagt	ctaagtgtgg	caggtggggg	gcaggagagt	126660
ctagggagag	aaagaggagt	aggcacccct	gccagctcct	gcagagttta	ccctcaaggc	126720
cggaaggaac	cctgatgcc	ggggaatggg	ccttgccctc	gagattgcac	atccttccct	126780
ctgtctctcc	tggggcagcg	gtcagtcagg	aggctggggg	aaagctctgt	aatcctccag	126840
gggctagcgg	ccatcagggc	tcacactctg	gtgagcttgt	ggataagggg	taggattaag	126900
ggatcagaga	aggatttggc	ttcttttggg	gtcaagtcc	tagggaagtg	gagatcagag	126960
ggtgactctg	acaggaaggg	aagtgccttg	gctgggcac	aagagacttt	tctggccctt	127020
tccctgccaa	cactttgctg	tgtgaccttg	ggtaagtgcg	ttgctctctc	tgagctccag	127080
tcatcacctc	agttagaactg	atgcttgaac	cagaggaatc	gaggggacct	ttgaggcttt	127140
gaaatctcca	gttctaagcc	ccaaacctca	accctcatga	aaccactca	gggtccccac	127200
tgtgcttcca	cactccacct	ctgctgtgtt	cagatgaggg	gtaagagaca	ttgctcctcc	127260
acccacagtg	ggtctaagaa	actcgggagg	agaaagtaat	cgtgaaacgc	cgacggggg	127320
aggggtgaga	agggccgaga	aacgcggagg	tgggtgtaac	gaatggaaca	gcagccgctg	127380
tgtcactgag	tattacatca	caccagcct	acacacgcac	ggggcccggc	gctcacacac	127440
acgcggagga	cagccagcac	gcaccgacgc	agcaccgacg	cagcgcaggg	aggggcccggg	127500
gacactcacg	gtggggccca	aaagcgagga	gcagcacact	gggagtgtgg	atcttccacc	127560
ccgcacctgt	gtgctccccc	ctctggagga	ggaacaccag	ggcagctggg	atgccagcgc	127620

cacactcggg	gcctgtcagt	cccatgcgtg	cacacctggc	tgagcagcac	tgcatttggg	127680
gagcacctgg	ctcacgccac	tacccaaaat	ca'agataca	tacacacatt	cacgcacacg	127740
gcaacctcag	gagcgtgaca	caacacacac	aaaaccacca	ctaagcaagt	gcaatttgcg	127800
gccttggaga	ccccacactc	aaaatcacca	accctcag	ctctcccagg	gtctctgaac	127860
cccaaggagc	cccaggatgt	cagagtgcag	aaacaagtct	tcctcccctc	tgccttcaaa	127920
agcctaggac	gttgcttgaa	gcagaaggtg	ttcagtcact	gtgtgcccag	ggaatgactg	127980
cctggccttg	gggggtgcagg	ctcccttttt	ccccaggcaa	aactgccaga	agaaaatccc	128040
aggagtacc	tggaaatcat	aagaaagtgt	agaggtcaag	ctagttccgg	cctagaactt	128100
tatcagctat	agtgcaggca	aaggccaggg	atgatgggag	gccctgcacc	cctattaaaa	128160
tatgagtaca	gacacctgca	ctccactctc	tagccccag	gctctctggg	cctgcttttc	128220
catcagtatc	ataataagga	tggatcatat	ccaaccttca	aaagttactt	tgggggaaaa	128280
aaaaaaaaaa	gctttggctg	gatgcggtag	cttatgcctg	aaatcccaac	actttgggag	128340
gccaaagtgg	gaggattgtt	tgaggccagg	agttttagac	cagactgagc	aacatagcaa	128400
gaccccatgc	ctacaatttt	tttttttttt	tttttttttt	tttgatacag	agtctcgctg	128460
tgtaaccag	gctggagtgc	agtgtgtcga	tctcggtcca	ctgcaagctc	cgccctcccg	128520
gttcacacca	ttatcgtgtc	tcagcctccc	aagtagctgg	gactacaggc	gcccgcacc	128580
atgcccggct	aaattttttt	tttgattttt	tagtagagac	ggggtttcac	cggtgttagc	128640
aggatggtct	cgatctcctg	acctcgtgat	ccaccgcct	cagactccca	aagtgtctgg	128700
attacaggcg	tgagccaccg	cgcccggcca	aaaattttta	aaaaattagc	tgggtgcagt	128760
ggcacggggc	tgtggtccca	gctcctcagg	aagctgaggc	aggaggattg	cttgagccca	128820
agtgatccaa	gctgcaataa	gctgtgatcg	taccactgca	ctccagcctg	ggcgatggag	128880
caagaccctg	tctccaaaag	aaaaaaagaa	agaagttttt	aagtaactgc	gaatgaggag	128940
agcctggggg	gtaaaatgca	gattcccagg	ctgtcccccc	aggaattctg	catagttcct	129000
aggactggct	gggtggcctca	cttagagacc	cgacccttaa	ggccccctcc	ggcacaaga	129060
ggctctgact	ctgcaagggc	gaaaagtaca	ggaaagtaag	ggcactgggc	accagtgggc	129120
tggaagacc	agaccccaga	gtgagtccat	ttcacacggg	cctcagatct	ccaaaggggtc	129180
ccaagttact	tccagtcatt	ctccaatggg	gtgactttgc	cccccagggg	acatttggca	129240
atgtctggag	acatttttgg	tgtcacaact	ggaggcaggg	tgctgctggc	atctagtggg	129300
tagaagacag	agatgctgct	aaatgcctta	tatagggctg	ccccacaac	gaggaactat	129360
ccggcccaac	tgtcaatact	gaggcagaga	aacctgacg	ttagtctttt	gacattaatc	129420
tctagacaag	gtcaaacatg	caatagttaa	aacaggaatg	aagagatgat	cattcttcaa	129480
ccaatttgcg	gtgctttcta	caatggcctt	ttggcattat	tttttaatat	atgagaagcc	129540
tcagaaagtg	gaagtggcca	ggccacttga	ggctataacg	ttgtcccctg	agccccaga	129600
catgggagca	ccagggtcct	aggcctttat	ttttattttc	tattttttcc	cctgaaacag	129660
ggtcttgggt	gtttggccag	gctggagtgc	aagggtgtga	tcgttgctca	ctacagcctc	129720
aaactcctgg	gttcaagcga	tctcctgccc	tcagcctccc	aagtagttgg	gactacaggc	129780
acatgccacc	atgcctggct	aatTTTTTTT	tttttcttgt	aaagacaggg	atctccctta	129840
tgttgcccag	gatagtctca	aactcctggc	ctcaagcaat	cctcctgcct	tggcctccca	129900
aagtgtctgg	attacagggt	tgagccacca	tattcagccg	ggtctaggcc	ttttaccaag	129960
ttggggggct	ggccccagc	tggcactcct	ggcctggaag	cccacctagt	aagttctgct	130020
tcccctcccc	acagctcccc	cctcggtcct	cctcctgctg	atgctcctgg	ccctgcccct	130080
ggcgcccccc	agctgcccc	tgtctgcac	ctgctactca	tccccgccca	ccgtgagctg	130140
ccaggccaac	aacttctcct	ctgtgccgct	gtccctgcc	cccagcactc	agcgactctt	130200
cctgcagaac	aacctcatcc	gcacgtgcg	gccaggcacc	tttgggtcca	acctgctcac	130260
cctgtggctg	ttctccaaca	acctctccac	catctaccgc	ggcactttcc	gccacttgca	130320
agccctggag	gagctggacc	tcggtgacaa	ccggcacctg	cgctcgctgg	agccccagac	130380
cttcaggggc	ctggagcggc	tcagtcgct	gcatttgtac	cgctgccagc	tcagcagcct	130440
gcccggcaac	atcttccgag	gcctggctag	cctgcagtac	ctctacctcc	aggagaacag	130500
cctgctccac	ctacagggtg	gcctgccctg	ccccaccctt	cagccccctt	ctggtttctt	130560
ctctctgtgg	gccccctctg	tccccgaccc	tggcgtgcgt	ccctcctctc	tccccaggcc	130620
acccttcctg	cctcagcatc	tccatttctc	tctgtctatg	tctcttttct	ctcttacatt	130680
ctccaggggc	tttacttttt	cccttctgcc	tctctacctg	tttaggtccc	ttgctgttcc	130740
tctctctctc	tctccctcta	actccacaac	cttcacctct	ctgcctctgc	ctgtctgtct	130800
gtctatccct	ttccatccat	cactgcctct	ctcactaact	tgccctcccc	atctgtcttc	130860
tgccctctct	gtctgtctcc	cttcacacac	ccactccgca	tacaccccc	tgtctgtctg	130920
cgtgtgtgta	tctgtctctt	tctgtgatct	caagtgtttg	ccttcagggc	actctgcctt	130980
ccccagggt	ccctgcccc	aaggcctttg	cagctgtttt	tctcaccac	cctcaagtct	131040
gccacatca	cggtgaagta	gagagagaag	gcagagccac	agccactggc	atccccagag	131100
aagttgcgct	tctctccaat	tactggggca	atgggacggg	agaagccac	acccttctta	131160
gattccatt	ttccaaacct	gtcatctcaa	tgagggggaa	gaaagaaaag	ggtaaatctc	131220
tggttatgcg	ctggagaatg	gatgctctga	aaatgggaag	aataccagta	attgttattc	131280
attgttatta	ttattgatct	aattattgtt	tattgtgttt	atgctgactg	tttgacacgc	131340
aaatcatccc	actccatttc	cccaggaagc	aatacacac	cctccaaacc	accctgagag	131400
aaaatcttcc	cttggctaca	gagcctccgg	ctggaagggg	gtgaaaatat	ccaaattctg	131460

ccctctccct	acttgaacct	ggaacgtgct	tcctctgcct	catccagggc	tagtgcctaa	131520
ctagtattca	atctgctagt	tggaaaatca	ggcagtgct	gatgatgcta	atgataataa	131580
caatagccat	aacaacctaa	caaacatact	gagcaccac	tacgagctag	atgctaagaa	131640
tacagtagtg	aacagaacag	accaaaccac	ctgaccttac	agagatacca	ttcccatgag	131700
gagggaaaaga	agtaaaatgc	acgggtatatt	ggaaaaatat	gtcttatatt	attcttattg	131760
ttgcctaaat	agtgcacgta	atagcagtag	ccgccaccac	ttagtgggta	cacagggtca	131820
gccacagtgc	caagcacttt	ataggtatcc	actctgccat	ttacaagcgt	gtgacatttt	131880
ttttttttac	ctcctcagac	ctcagttttc	tcatctgtac	aatggggtag	caagagcacc	131940
catctcctag	ggattttgaa	agcattaaat	gcatgaataa	tttgtaaagc	acttagaata	132000
gtgcttgcca	tacggtaagt	gctatataaa	tgcttggttaa	aatactattt	taaaaaaaga	132060
aacgagcctt	atttaacatt	ggtttcagtg	aagtggccca	acttggaactc	catcctgaag	132120
atgtgggtca	acttcaagga	ttatactaag	gtcatgagtg	agtcccagaa	attgcacctc	132180
acagtttatg	aagtgcactc	agccacctca	tctcatattct	acagcccagt	tgggagatta	132240
ttttcacctc	cttggttaaca	atggagaagc	tgaggctggg	ggccctgaag	accctataga	132300
gatatagtca	cctccaatca	taaatctttt	caacctttgt	cggtgtgacc	ggaggcttat	132360
gtcttctcac	catcatgttg	agcctcacaa	caacctgggtg	atagggacag	ttaggggcac	132420
tagggacatg	gaatgaatgt	tcctgaggcc	acacacccag	gaagagctgg	cgcttgaaac	132480
tcatggtctg	gctacaaggg	gacagtactc	tggagtacaa	ttgagcaggg	tcatttttga	132540
aagcacacag	tttggtactca	gcaagacctc	ggttcaaatac	ctggctccta	tatatatgac	132600
tttggaacaaa	ttacttaacc	tctctcagtc	tccatttctc	catctctaaa	atggcaatca	132660
ggatagtact	taataataat	cttttttttt	tgaagcagcg	tccactctta	tcgccaggc	132720
tggagtgcag	tagtgcgac	tcggctcact	gcaacctctg	cctcccaggc	tcaagtgatt	132780
ttcctgcctc	agcctcctga	gtaactaaga	ttacaggcat	gtgtcactac	acccagctat	132840
tttttgtatt	tttagtagag	aagggtttca	ccatgttggtc	caggctgggtc	ttgaactcct	132900
gacctcaggt	gatccactca	cctcggcctc	ccaaagtgtc	gggattacag	gtgtgagcca	132960
ccatgcccag	ccaataataa	tccttattta	agaagttttg	taaggattaa	aatgtaaggc	133020
athtagcaca	aggattaaaa	tgtaaaggcat	ttagcacata	tgggcactat	aataataatt	133080
actactacta	ctactactaa	tactgagatc	aaatactact	acaaattgat	catgcattta	133140
atgctttcaa	aatctcctta	tcaatatata	ttagtatttt	aggaggaatt	tggagtcaga	133200
gggctgagc	ttgaatcccc	gatctactat	tttctgactt	atttaacttt	aagcaggttg	133260
ctaaccctct	ctgaacctca	cttactttat	ctgcaaaactg	ggaataatga	aaataatacc	133320
ttccaccaag	aatggctgta	aataggaaac	gagttagtgt	atagaaagcc	catagttcag	133380
gctgggtgtg	tggcccatgt	ctgcaatccc	agcaacttcgg	gaggccaagg	tgggtggatc	133440
acttgaagtc	aggagttcga	gaccagcctg	gccaatatgg	tgaaccctg	tctctactaa	133500
aaatacaaaa	attaggcagg	cggtgtggtg	ggtgtctgta	atcccagcca	ctagggaggc	133560
taaggcagga	gaatcacttg	aacctgggag	gtggagggtg	cagttagctg	agatcgtgct	133620
actatactcc	agcctgggtg	acagagcaag	actctgtctc	aaaaaaaggaa	aaaaaaaaaaa	133680
aaaagcccat	agttcagtg	tgaagaaatc	atgtttattat	gaccccatcc	tccattgact	133740
ctcaggccaa	caacagcaat	caggacctga	ggtcagcaaa	ggcttgggca	gaggggacct	133800
caggtggaca	ttgggtgctt	ctgaaatggg	aagtgtttgt	tctctacgcc	cctggcatga	133860
atggtaccag	gcatcatggg	aaggaaagcaa	cttcacacct	ggccttttat	agaggagatg	133920
gaaaaacacag	cctctgcctg	tgaactgcct	ggtagggtcg	ggctgggaga	tgccacaggc	133980
aggtgaggaa	acatgggctg	gggtgagatc	cgcagggtgc	aggtgtgacc	caagatggag	134040
ccaggcctgc	cccaaagggg	agctttggag	gaaactccac	cagaggacca	cagcttttca	134100
gaatggggaa	gggccaggca	ctgtgccagg	tgagttcatt	catcaacaga	tatttactga	134160
gtatctacca	catgccaggc	aatgttccag	gtgccaggga	ttcaggagag	aacagaaaca	134220
gtggccctgt	tctccagag	catattccct	actcaagtgt	agccagatga	taaagacact	134280
tgttttcttt	cttttttttt	tttgagacga	agtctcgctc	tcttgctcag	gctgagtg	134340
agtggcacga	tctcggtcca	ctgcaacctc	tgctcccag	gttcaagcga	ttctcctgcc	134400
tcagcctccc	aagtagctgg	gattacaggc	atgtgctacc	atgctgggtc	aattttttgta	134460
tttttagtag	agacggggtt	tcaccatgtc	ggccaggctg	gtcctgaaat	cctgaccaca	134520
ggtgatctgc	ccaccttggtc	ctcccaaagt	gttggtattac	aggtgtgagc	caccgcaccc	134580
gccgacactt	gttttctctt	tcagtcatca	cagtggcctg	catggttttt	gtttgttttg	134640
ttttgttttg	tttttgtttt	tgagacagtc	tcactatgtc	acccagctgg	agtgcatggg	134700
cgcaatcttg	gctcactgca	gcctcacctc	ctggggtcaa	acaattcccc	catcttagcc	134760
tccccagtag	ctggaactac	agacatgtgc	caccatgtcc	agctaatttt	tctattttat	134820
agagacgggg	tttcaccatg	ttgccaggc	tggtctcaaa	ctcctgaact	taagcaatcc	134880
acccgctctg	gcctcccaaa	gtgctgggat	tacaggcatg	agccaccgta	cacagctggc	134940
ctgaatgggt	taaaaatagt	ctttatgctc	aagcagatca	gatctcagtt	tgaattccag	135000
ccacacctct	aatttgctct	atggctttgt	gcaagttatt	taaccactct	gagcctcgat	135060
ggaccatctc	gtgaaatggg	gataacctgt	accttgccga	gcaggggttg	tgaggattaa	135120
aggagatact	actgagctca	cagcccaatg	tctggtacaa	agtgagtatc	caatgaatgg	135180
tagctatcca	ttaacaccag	ggaggacacc	aactgaagct	cagcaaaaata	aaagcacagt	135240
ccaaggtcac	ccagctagta	aggaacatga	cctagaattg	gcccaggtct	gtctgactcc	135300

agagtgcagt	tgttcagagg	tctctggagt	tggaagccac	gttccactgc	atattagctg	135360
ttggacccta	ggcgagtcac	ttcacttctc	tgaggctcca	tctcgtaatc	tctgaaatgg	135420
agataataat	agtatccacc	tcataggggt	gtgacaatta	agttactata	taggatctgt	135480
gtagcacaga	gcttggcaca	tggttaagagc	tcaatcagtt	acctgcttga	caatgctgac	135540
gccgatgatg	acgatgtatac	ccatcctaga	ctgatgagct	ctgtaagcgg	gggtgcctgg	135600
cacagagtag	acactcggta	cagctctgtg	gaatgaatga	ggcacatccc	agaactcacc	135660
aattcataaa	aatcagatgc	agatgggatc	ttaaagatca	cctatcctaa	gtcccttggt	135720
tcacagatga	aaagacccag	gccagagag	gtgcttgagg	ctgcgcaagg	tcacacagcc	135780
aagcagctca	tttgattagt	gtcagagcca	agagctggga	gtttggaggg	aggcaagggt	135840
agaacagga	tgctgtcagg	gaagcaggca	gggatgctgt	gttaagattc	caaagtggatg	135900
cagagagctg	tgaaccggcc	agtggggagg	caagggaat	gtggtttttg	aaatggaaga	135960
ggatgacttt	agcagaggct	ctcagcccag	agggagggga	gatagggagg	ggagataggg	136020
aggggcgggg	ggagggctag	ggctgtgaaa	gtcaagagct	tattaatgca	tagagaacgg	136080
ttttaacagt	ggagagagga	aggaccggat	ttgaaagcta	cattcaagga	agtggcaacg	136140
ggatttgcca	acagcttgga	tggggggagg	aggcaatgga	ccccaaggca	gaggctcaga	136200
gaagggaagg	gcaggacttt	ttgcagagaa	acaaaaggga	aggagaggag	gttagaatca	136260
agaaattctg	tgggccaaaa	cctggggctg	tggggtcaaa	gcacctgaat	tccttaggat	136320
ctctggaact	ttggtctact	cttctgacct	cccgaggtcc	cccaaaatgt	ggattacccc	136380
tgctcactct	cccccaaccc	ccggccccct	atcgatcctc	tgaccataca	tctctgggtg	136440
tgctcactct	ttgctgacac	ttcataaaaa	gaggaacccc	atttaggtgt	tttgagtggc	136500
agggattcca	agctacccc	ctggatgggc	ctggaagaga	acaagagcac	caggccatgg	136560
tgagtcaagg	tgaggccaag	gaggtgcaag	gagccagctg	gaggcctgag	ccaggatttg	136620
gggtggtggc	agcagggggc	ggagaatggt	ggtgtcagag	gcagccgaga	aggttgaggg	136680
ggacggatct	caatgtggcc	aagaggaggg	ctcttgccac	gctcagttcc	tgtagcgaag	136740
agggcggaag	ccagatggga	gggggcgaga	acaggcagga	gcacaggaag	gtggaggctg	136800
tggtgtgagg	ctgggagtca	atgccctccc	ccaacctgag	gcctccgacc	aggtcctgg	136860
gtggcaggca	tggggaggaa	agcgtctccc	caggcagtga	gggagggaga	gccacagtca	136920
gggaacaggc	cccctgggtg	aactggcctg	agcagagtgg	atgctcctgt	tctgagaccc	136980
agacctcctg	gaacctgctg	accacagtga	tgccctgcac	aagaggggag	gacctcaagg	137040
cagtgaggtc	agggagctga	agtcctgctt	ccctctctgg	caagccctta	tctctttgag	137100
ccccagtgct	ctcctctaaa	aaagttagct	gggtctgatg	gtgccaaagg	attagctccc	137160
aaagtcagctg	atcatcagaa	tcccctgggtg	agctggttat	aatgcagagt	ccaggaatcc	137220
ccactggccg	tgggccacac	acacccgctg	cccccgctg	ttaattctga	accatagttc	137280
caaggtcctt	tctgactaa	tgtggcctga	ttaggtgact	ccctagcacc	aggcaggtgg	137340
gacagcgctt	ctaaggggag	tagtaatgca	atgtggcttc	cttcctctcc	ccccctgccg	137400
cctctggggg	tgagagctgat	gcccctcacc	ccaataccca	gcctagtagc	agtacttttg	137460
ttcccccagg	gagctcctct	tttaaagaaa	agggacagga	ccaattgtt	actgagcccc	137520
tattgtcata	gtagccacca	tttattgatg	gttgactatg	cacctgccag	atactgtacc	137580
cttaacagca	tttatcatcc	aacctcctt	tagcctgctg	aggggggttat	acataataag	137640
gaatattgta	catactgagg	aacctgagac	tccatgaggt	taaaacttgc	ctaaaataac	137700
acagctaggg	aaaaggcaag	ctggattttg	aactagggct	ctaagtgtcg	agcctgtggg	137760
cttcataatt	ggaccaaatc	cctgtgtgct	gggcacgtgt	ccagcacttc	cctcatatga	137820
tctttatgtg	aacctcctc	tggaatcctc	agaacaaacc	caggaagttag	gtatactcat	137880
ccccatttta	cagatgagga	aacaggcaca	gagagatgac	tggttggtgg	aagttaagaa	137940
taatggctaa	caaaacaaaa	caaaaacaaa	aattaaaaaa	aaaaaaagaa	taatggctaa	138000
ctcatggaac	tcatagaact	ccacaaggaa	aggtgttcta	agcaccttca	tacatgctgc	138060
ttcattttaat	ctctacatta	tacagatgag	gaaactgagt	cacagatatc	ctgagtgtgct	138120
tgcccacggg	ggcatcagtt	aatgacagat	ccaagatttg	aaatcagaaa	ggctggctcc	138180
ccagtctcca	tacttcacca	aaccagaagt	tctgaaactc	aaactgtggt	cctgccaatg	138240
gccacactgg	cttccctggg	gaacctgtag	acatggggat	tcccaggctc	caccccaaac	138300
ctcctgaatt	agaaactctg	ccccccgccc	caccccgctc	agagatccgc	aggggatcct	138360
aatacacccg	aaagtttagg	aacctactgac	ctcaccaata	ccactttttc	cacagcaaat	138420
aggttagagg	aggcagaatc	caaatccagg	atgctatgaa	tcaaaagggtc	aaccttttct	138480
cttctgccac	ggtgcacccc	cttccctccc	ccggccaagg	ccccagcggg	gtctgcaccc	138540
gtcctcaggc	ccattctctt	cttctgtgcc	ccactccacc	ccaccagga	tgacttggtc	138600
gcgacactgg	ccaacctgag	ccacctcttc	ctccacggga	accgcctgcg	gctgctcaca	138660
gagcacgtgt	ttcgcgccct	gggcagcctg	gaccggctgc	tgctgcacgg	gaaccggctg	138720
cagggcgctg	accgcgcggc	cttccgcggc	ctcagccgcc	tcaccatcct	ctacctgttc	138780
aacaacagcc	tgccctcgct	gcccggcgag	gcgctcgccg	acctgccctc	gctcgagttc	138840
ctgcggctca	agcctaacc	ctgggcgtgc	gactgcgcgc	cgcggccgct	ctgggcctgg	138900
ttccagcgcg	cgcgcgtgtc	cagctccgac	gtgacctgcg	ccaccccccc	ggagcgccag	138960
ggccgagacc	tgccgcgcgt	ccgcgaggcc	gacttccagg	cgtgtccgcc	cgcggcaccc	139020
acgcggccgg	gcagccgcgc	ccgcggcaac	agctcctcca	accacctgta	cggggtggcc	139080
gaggccgggg	cgccccagc	cgatccctcc	accctctacc	gagatctgcc	tgccgaagac	139140

tcgcgggggc	gccagggcgg	ggacgcgcct	actgaggacg	actactgggg	gggctacggg	139200
ggtgaggacc	agcgagggga	gcagatgtgc	cccggcgctg	cctgccaggc	gcccccgac	139260
tcccagggcc	ctgcgctctc	ggccgggctc	cccagccctc	tgctttgcct	cctgctcctg	139320
gtgccccacc	acctctgact	gcggtgctga	gatcgaagag	gccagtgtcc	gatccccgct	139380
tcccgctcac	ccggggctgc	ggctccggcc	ccagtccgcc	caccttccct	ggccttgcctg	139440
cctccctttc	ccctcccagc	tcctctcctc	cccggggagc	aggccgcctc	tccttgccctg	139500
ccccctgggc	tgtcctgact	tgtggcagcc	ccaagagggc	gtgtgtgggtg	gctcagccct	139560
gccctcccga	gttctggcca	ttactcttc	cccatcccaa	ggctgggggtg	gggcccccca	139620
ggcagccgct	gacccgcact	cctaagggcc	cacagcggac	accagagggg	cttttgcctg	139680
cagagcgctc	tccaccagca	gagcctttgg	aagctccccc	agggagcccc	accaggacc	139740
ctttggggga	tgcctcagtc	agggccaggc	tgacctgac	ccctgcttac	cctagtcccc	139800
tcaacctcct	gacactggag	gaatactttt	ctcctaagtc	tacctggac	acttttttagg	139860
gcacctggag	agaactttcc	tctccactgt	ggcccctgcg	tggtgaagat	caaaagaagt	139920
tgtttgggaa	aaaaaattta	ttaaaaaatt	ctattatttt	atctactgta	agatttgttg	139980
acttgggacc	ccgaaagcgg	gatgaggtct	cagaatgtaa	ggattgcagg	gccaggaggg	140040
ttggagaagg	ggagccgtcc	cccgccatca	aagagcttcc	tggtggctgg	aggtgggtgtg	140100
cgctcccccg	ccatgaggag	gagctgaagc	cctgcattct	aggtgaggcg	cagtgtggca	140160
gccaagagtg	ggtgctgggtg	gcacctcttc	tcttcatttg	tccaggggaa	gagctgcagc	140220
caaccctgag	tggtctggcg	cctgaggaac	taagcctggg	gaagacctgc	tgtctggtta	140280
acagccctct	tccagaccct	gttccttcag	gaaacaagag	cagttctcct	gcaaggagga	140340
gtcacataca	cactcctggt	cacagacagc	cccaacatgg	ctttgggtaa	atgtgaacaa	140400
ggcactgctc	cctcagggaa	acacagcccc	atgccagagc	aaacacctta	gcaaacagag	140460
accaaggctg	ggtttcccg	tacacttgcc	tccttggtca	agtgccttg	tgcagtgcac	140520
agcgtacaca	cctgcacaca	gcaaccctgt	gggtatgtgg	tctctctctc	agctcctgtg	140580
aggtagaagc	catcagggat	gaaccaggtc	agagaagcag	gtttccaaac	aggctagaag	140640
agggaccgag	gaactcgggt	gatcagaggg	acaggaatcc	caaattggga	tgcattactg	140700
gcttgaggta	caatcagaac	cttcactctt	ctggtgtgtg	gaagagaggc	tggggactgg	140760
gaagagctca	ggctaagaag	gacttgggtt	gggatttagg	ggtgagtctc	atcagactga	140820
gcacttggag	agaagtgttg	tagtttgaat	ttggagctaa	gaatctagct	tgggcagggt	140880
gtggctcgct	gcacctgtaa	tcccagctaa	ttgggaggct	gacgtgggag	gatcacttga	140940
ggccaagaat	ttgagactag	cctggacaac	atatcgagac	tgagtctctt	aaaaatgttt	141000
ttttaagaat	ctagtttgga	gtgggggtgtg	atgtctcaac	gtctgtaatc	ccagcactct	141060
gggaggctga	ggtggacaga	tcacttgagg	tcaggagttc	aagaccagcc	tggccaacat	141120
ggcagaaacc	ccgtctctac	taaaaattca	aaaaaattag	ccaggcgtga	cggcggtgct	141180
ctatagtccc	aggtaactcag	gaggctgagg	cacaagaatc	actccagcct	gggtgacaga	141240
gactctgtct	aaaaaaaaaa	aaaatctagc	ttgggagggtg	ggaatagaaa	gatagagggg	141300
gcctagatgc	tagggcttga	ggaagcaggc	tgagggttctg	tgattctggc	tagggagggtc	141360
aaatgatctt	gagaagaaga	gaagaaagga	gaagaaatca	gcatctaagc	ctgaggcagg	141420
tagactccgg	ttaagggtgt	ggggtgggct	gggggagagt	gagagcagct	ggtcagaaac	141480
ccaggagctc	cggagtctgg	ggtcttgcat	gggcttgtgt	caggctggct	gtgaggaggt	141540
taatgggttg	gattggaggg	acagccagac	aagagctctg	gtggaggagg	ggctgctggg	141600
gcctgggag	ggggagggga	gctgctggta	aattagaggc	aggctgtcca	ggtcatagaa	141660
ttatcattgt	gaaatattca	tgggccatcg	gtccagatgc	tatttcagaa	cagtgaagc	141720
aaggaggatg	tgtgagcctc	aggaagaagc	ctgaagcaaa	gcaactctcc	accaaccccc	141780
accctcccca	ccaccagccc	agacagaccc	acggagcggc	atcacgtgca	caccacact	141840
cccgagctct	cacacacact	cgcaccaagc	agagccatgt	agcacgtgca	agcacaccaa	141900
ccaccacagg	gtcccacaaa	caggcagggtg	tcccctaaat	tctgacatgc	acactgacat	141960
gcacacccac	tcaatcagga	cccagcagag	atcacctcca	gcgatctcac	atgcgcagac	142020
ccccaaactc	tccaaacaac	ccgatttcac	caccttgacc	cacacaccct	gagataggag	142080
ggatgttcaa	ggccatccag	cccaaccccc	accaatgctc	tgatggggaa	actgaggcca	142140
tagaaaggaa	gggatttgtc	tgagattcct	ctatcccctg	aaaaaagcaa	aattcattca	142200
cctccacat	tctgagtgtg	ccccattctc	gcattttcgt	ctgccagaca	cccagcctag	142260
ttgtaattaa	ctcctccctt	tctctaattt	cctgcatcta	ttcagttacc	cagttccccca	142320
cccagccaca	gtctatccct	tccttcccat	tctccccacc	acctccctgc	tccagctact	142380
cattacctca	tgcttggaa	ataaaagaaa	actgcgataa	cctcctcgct	ggtttcctac	142440
atggaatctc	tcctccctc	ccaccagcc	ataccgtggt	gaccagattc	atctgatcaa	142500
aatttgcata	tgttatgatg	tactcagga	gctgtaatg	gcttccta	gcctataggg	142560
taaaggtaaa	acaccttagc	agagcatcaa	agatccctca	gagtctggta	ccaactgctt	142620
ttctagcctt	ttctctcaca	atctcatccc	aaaccttcac	tccagctaga	acgtttgtat	142680
catactggcc	accagttatc	atgtatgtga	aaccaccaa	ccgactttga	gtgccccct	142740
aaaatttctc	agtctctcct	gaagtaggaa	acctcttccc	cctcctcaga	tctcagactc	142800
cagagccctt	tccaaggcc	aagactgcac	ctctctgacc	atatacaggg	gttcttcaaa	142860
gcagcagaca	gaggctcagg	ctctggctcc	ctccaagcag	acggctgccc	ccgactggcc	142920
accttgggaa	gcacagccag	gtttcagtcg	tctagaacag	agaatgagca	tctaaccgcc	142980

tggggagagg	actaggacac	cagatgataa	ggtttataag	cccttaagcc	tctaaggttc	143040
ttacacccag	agtagggggg	ggacggttct	cagccctggt	tccctagctg	cgggctccca	143100
atthttcgatc	cctaataccga	gaggaactcc	tctccaatga	aatacagact	tgggactctc	143160
aggacactgt	ggaagggaag	tttcccaaca	gactctgaga	gtccaggagg	ccagggatag	143220
accagggtggc	aggcccaagg	tccagctggg	gtcaggtttc	tatatgaatt	tttaatgctt	143280
ccagatagac	ttgtcagatg	ttctgaaaac	tgagcatctc	ctttcacctc	tgtacatgat	143340
gcccttctcc	aacccattg	cccctgcagg	agggcaggcc	tgggacagat	attcagtggc	143400
ctctggagaa	acggttttgg	gacagtagaa	gggtaaatga	cctagttagt	ttcccactag	143460
taagctgtgt	gaccttgggc	aagttactta	acctctctga	acattagagt	tctgtgggtt	143520
tgthttttgtt	ttgtaagctg	gggacaatag	tgccagccta	aatcaatttg	ttgtggggac	143580
tcagtgaat	agcccatggc	aaagtgcct	acatgcttgc	tgthattatt	ctctttcctc	143640
aagttctgcc	tccctcttcc	agcttttctt	ccaaccccaa	agatgtctct	ggctattgct	143700
tcgaaggtag	gaacttttgt	tggttctccc	ctttctcttc	aggcccaaac	tccccacctc	143760
aagatccttt	ggccttttga	gaaacttcag	gtgaggaggt	ggcagagaaa	taagaaagtg	143820
tgcaaggctg	gtggagttag	agaggaggat	agatggcgaa	gccctagcag	agggggagga	143880
agtgggcagt	ggagagagg					143899

<210> 16

<211> 215980

<212> DNA

<213> Mus sp.

<220>

<221> modified base

<222> (1001)..(1100)

<223> a, t, c, g, other or unknown

<220>

<221> modified base

<222> (2123)..(2222)

<223> a, t, c, g, other or unknown

<220>

<221> modified base

<222> (3728)..(3827)

<223> a, t, c, g, other or unknown

<220>

<221> modified base

<222> (5168)..(5267)

<223> a, t, c, g, other or unknown

<220>

<221> modified base

<222> (7481)..(7580)

<223> a, t, c, g, other or unknown

<220>

<221> modified base

<222> (8849)..(8948)

<223> a, t, c, g, other or unknown

<220>

<221> modified base

<222> (10375)..(10474)

<223> a, t, c, g, other or unknown

<220>

<221> modified base

<222> (12270)..(12369)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base
<222> (13438)..(13537)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (15902)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (15939)..(16038)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (18223)..(18322)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (20974)..(21073)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (24403)..(24502)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (27574)..(27673)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (30892)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (30901)..(31000)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (34443)..(34542)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (38205)..(38304)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42373)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42386)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42393)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42461)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (44809)..(44908)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (51380)..(51479)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (56740)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (56765)..(56864)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (62818)..(62917)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (68518)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (68534)..(68633)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (74552)..(74651)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (81446)..(81545)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (88519)..(88618)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (93791)
<223> a, t, c, g, other or unknown.

<220>
<221> modified_base
<222> (93794)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96565)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96570)..(96573)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96579)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96590)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96596)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96602)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96616)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96629)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96633)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96668)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96715)..(96814)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (104447)..(104546)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (114521)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (114527)..(114626)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (127063)..(127162)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (139133)..(139232)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (151051)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (153242)..(153341)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (164706)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (164708)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (164710)..(164809)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (182242)..(182341)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (192158)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (192192)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (198842)..(198941)
<223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (199437)..(199438)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (208276)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (215974)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (215976)..(215977)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (215979)
 <223> a, t, c, g, other or unknown

<400> 16
 ttggggggtat aaaccagaa gtgggattac tgcaccatac aataatcctc taacttcaag 60
 caattttttcc acaatgggtg tatcatttta cattccact ggctacgaga agggttccca 120
 cttctacaca tcttcaccac catttctggt tttgtttttg agtaacagct gcctaagac 180
 tgtgaagtgg tatcttatct cagtgttgat ttgcatttct ctgatcatta atgtgggaag 240
 gcatcggttc atatgtttat tggctgtttg tgtatcatct tctttggcga tgttgattca 300
 agttatttgc ttgttttttt aattggagtt ttaaaaaatt gttgttgagt tgtgggagtt 360
 cttcattagc tctgcatatt aataccctga tgaaaatgat taacaagtat ttgcttccat 420
 tttggggggt tccattctgg gctgttttta ttcttttgat actcttttga ttctcaacag 480
 tttaatctga ctaaaattca gtttatttct tcttttaatg gccatgctat tgacacatcc 540
 cgtaatcact gccaaatcca gtcatagaaga gtttctttca agagatttat agtttttagct 600
 ctttaagttt gtcattgtctg ttctacttaa ttttgtatag tgtacaaaag tctaacttca 660
 ttcttttcta tatggcttgc tactagtata cgaagagcta aatttctctt tccttgagtc 720
 tcaacctctg atgtgtagca atttcttcag aggaaaacat ggtgggaagt tccttaaaaca 780
 taggatgctc catggaggtg aaatagttca tcctacaggg aagcttggtta aacacaggaa 840
 gtacatactc agcagctcta gtaagtgaag gaaactgact ggaggcacta ggtccctcct 900
 tccctacgca tatagaagct gtaaggattg ggaagagata ctgtcagggtc agctcagctg 960
 ctgcccggaa gaagctcaga cccactggcc tggctccaag nnnnnnnnnn nnnnnnnnnn 1020
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1080
 nnnnnnnnnn nnnnnnnnnn atcactcttt actcaggcca cctacacgct gtttatagcc 1140
 tggcctttgtc tctttggcta tacttctctg ttatgtctat gcctcccctc tttcttttct 1200
 tttctcttct cttctcatct catctcatct ttcttcaggg gggagcctgg tctagaactc 1260
 acaaagattt gactgtctct gtctccttgc actaattaaa aaatctttta caagcatctt 1320
 ttagcaattc ttacaggga attttggaa gttaaactct gattgttagc gggctgaaga 1380
 taacaatagc tctgatgata aattgcttgc caggcaagtg tgaaaatctg agtttgatcc 1440
 aaaaagccgg gtacagaggc caaagagtc ataactctag taggggcagg aatcagggat 1500
 ggggtgggtcc ctgggggttc ctggtttgtc agcgtagccc aattgggaat agccagggtt 1560
 cagtgaacga tgctttctgc aagctgagag aggtccttgt tcaatctctg tgacccaact 1620
 ggaggggagaa gagagccagc tctccagaag tggctcctct aactttgtgc atgcatgtcc 1680
 atgttcacac aggggaatga taatgcttaa aaggaagacc ggagggggtg tggtaatgca 1740
 cctccttttg tgacatgctt tcctcttggt catgtctgct cagggtgtgt cggcagcacc 1800
 aaaaaccagg tgtatgtttg taatcccagt attctctggt cgtcagtagg aaatgaaaag 1860
 cgagggtcatc ttctgtataga gttagcaaac tctaagccag cctcggtctac atgagacttt 1920
 gtctcaaaac aaaggaaaaa tcaaggagga cggctcccga gcaactgtcac ctgaagctga 1980
 cctctggcct ccacatgcat gtgcgcaaac acatgtcctg cacaacacac cagacaccgg 2040
 catctgtctc ccgacaaaag aacctgaaac cagtataact tgagaatttc ccattcatag 2100
 ttaccattgt gtgttctctg tgnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2160
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2220
 nnggtggtgc ctttctctta cccagtctag aagggtctga ggcagggtgg atggggcact 2280

ttgaactccc	acctaggcaa	aaaacccagt	gatctctg	ccagtgtgtt	gtttgcaagg	2340
gaataaggta	gagagccg	gaggaagaca	ttgggggttc	tatgagtatg	tgaaggggtg	2400
cacacaccac	acacacacat	tttttttgtt	ttaaattttac	aaacattaaa	ataggctgta	2460
atgtggctca	gtgggtagaa	aaacctgctg	tctaagcctg	gtacgagttc	aatccctgac	2520
aagctggaag	gagacaacca	accacaactg	ctagcagcca	gaagcactgc	ttgctaacac	2580
tcaagagagc	ctggagtggg	agacactgga	tccccagcag	gcaagcctgc	aagaagatgt	2640
gccttgcccta	gacaacggca	gaacaaacat	caaggctggc	agagctgtcc	aggactgttc	2700
atattaatca	tgtatagata	agagggaatg	gcacagacag	aacaattcaa	cacacggggg	2760
atgaaaggaa	aggaacaagg	cacacaaagg	acaaagaacc	tagcatacaa	gaaagcctaa	2820
gcagagagtg	gcacttccca	gaaaggagtc	ataaaataga	ctgaattcat	taaaacaaga	2880
gccaaagata	aacggctcaa	aaaactcacg	gaaaacaggt	caaaaataacg	tcacccatct	2940
gacagttagt	actgtcaact	taaccgtatc	tagaactcca	gcaggcacat	ctccaggcat	3000
gcccctgaag	gggtctttgg	actaggttaa	ctgacgtggg	agtgaaccca	tctatggacc	3060
gaagcctcag	acagaataaa	aaggagccag	tgagctgagc	gtcagtgtct	attgtctctg	3120
gcttcctgtc	tgtggctgca	gcgagacacg	gtgcttcctg	ctttagctgc	catgacagac	3180
cacaccctca	aaccgtgaac	caaaataaac	tcctctctac	attgctttta	ccaggcattt	3240
ggtcacacca	atgagaaagg	ttaactaata	cagcactcaa	tacttaaaaa	cataaacacc	3300
aaccttgctt	gcatgtgtga	gactttgaag	ctacggggcc	agttatgccc	aatgccaggt	3360
ctgctggcta	agggtagagag	tgcacaccta	taatcccagc	tgctgtggaa	tcagcaaaag	3420
cgctacagat	ggaaggcagc	cagggcagct	gagactgact	caaaactgata	gaggtgggag	3480
gcatagagaa	aaccagatta	atagagtgtt	ccccactatg	caagaagccc	tgggtttcag	3540
gacgagagaa	ctaagaatac	agaagtctac	tgtgtagaag	cactgctagg	tcacacagaa	3600
acatcactca	agtgtctctg	gatgtctacac	ggagggcgctg	tgaagtattg	cttcctgatg	3660
atctgtatct	actacagcac	tgctgtttta	gtatgcgctc	ctccactaca	gctcctcacc	3720
acaccaannn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3840
aaacacacac	caccagttag	agaaagttaa	tcaggccgaa	tggcggtttt	cccctgtatc	3900
caggtaccg	tcaggacggc	tcactgccac	tggcaactct	gcctgaacaa	agcccagc	3960
caacgtgggc	ttcaggggct	ctaaacactg	caatcaaagg	ttgtgtgtgg	gggtgggggt	4020
gctgtgcta	tcaaggatt	cccaaagcct	agatgtattc	atcatactca	caggaaagcg	4080
tgttcaacc	atcactcatg	agcagtcggt	accggggtga	cctattccct	gtagaaatgg	4140
gacggatgtt	ctggaaaagt	tgacagaaaa	gttgattcat	taggcaggct	ctttgcccaa	4200
gccctgagg	taagcaaaag	taactggcag	gagactaggt	ttgccattaa	tctgagacaa	4260
gatgaaccac	ttgcccatcc	tcctgacacc	taataactaa	tgaaagaaca	atggattgag	4320
ctggcattat	taaaaacgat	agaaacagaa	gtatcaatag	tcattgtgtc	tttctcccat	4380
atgtcaaaac	aatgtgtaag	atggcatcga	acacatgcag	aaactgttta	gggaacatgc	4440
tgaaaatatg	aagtaaaatt	aaaattggaa	agaaagacaa	tttgccctaaa	gcagctcaga	4500
gctggagaag	ggaccgaggc	agagataaca	gcaacgtgtg	gacatacgga	tctggggcag	4560
agcagtcacg	gactcagccg	gaaagggtgg	ggcagcctct	gaaggaggtt	aaggtaaata	4620
gagccacaag	gtgattggcc	caggagtggg	gccaccttca	cctcctgcct	caaaagtctga	4680
aggaatgatc	ctggagtctc	ccatctattg	atatatgaaa	ttcacagtat	gttttagaac	4740
ccactgaatg	atgggtagat	taactaaaag	aaatttaagc	ggggtggtgc	aggtccttta	4800
atcccagcac	ttgggaggca	gaggcaggtg	gatctctgtg	agttcgaggc	cagcctggtt	4860
ccaggacagc	cagagataca	tagagaaacc	ctgactcgaa	aaaacaaaat	taaaagctca	4920
tcaaaaaca	aacaacaaca	aaaaaaca	aaaaaaca	caaaaaca	aaacacccta	4980
tagtacctgt	tggtgagttt	gagtgtgtga	gtgagtgtgt	gttagagaga	ggggcgggga	5040
aagtgtgttc	tggaaatggg	agaaagagaa	tgtgcatgtg	tgtttctggg	atgtagacaa	5100
aactacatgt	cttccatcaa	atgcaatgtt	taattatcta	tgagtgtgac	catcttcatt	5160
ctgctaannn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	5220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	5280
caaaccagta	aacaaaatcc	tgtaagataa	agcctaagac	aagacacttc	ctggggctgg	5340
ggagtgtgct	agaccataag	gagttcataa	tccaggcggtg	agagcccag	ttcagggtccc	5400
tgggcttcca	agtcaggagg	agaccaagga	atcaacaagt	ctcgactttg	gtctctagtc	5460
ccatgcacac	acacgcgtgt	aaatacgtag	atgttcaactc	acacacagaa	gactgcacct	5520
ggctctctca	catctcagcc	aacatataaa	gcctgcatta	tcagaacatt	ctaggttcta	5580
gtttcagtc	actcttacac	agaatggcca	tcatactccg	tctacaactt	ctcctgatct	5640
accacgtgt	cattgtctca	aaccagaaat	aaccagctgc	aaccagctgc	gtagatcctc	5700
cctgatgcc	cagtcatgtt	cttactgaga	ctactaagtc	acaaggtagc	actctggatc	5760
caaaaagcaa	tatccaattg	agagttacaa	cctataagga	ggagttagcc	ttcattatag	5820
ggcactggat	tcccaatctt	taatccaacg	tcttcagcag	atttcataac	ttccaagtcc	5880
atcaaaaaca	ctactttcct	acaaagacag	acacaagtta	gaattaaaga	ctctgcagcc	5940
tttcagatga	gttactaaga	agcttacttt	agtagttgtc	tggctaaaac	tgtatccttt	6000
accaaccttt	tctcattctg	gactaacttg	agaagtatta	attcctaagt	aaatacttca	6060
cttattcttt	ccccacatct	ccaatgtttt	tgtctttaat	ttattatagg	gcaattcatt	6120

tcctatctag	ttccctgatt	aaaacagtag	accttgctgc	atgccattat	cctcatggag	6180
gcaactgatac	aatttagatt	attaaatata	aaaccctaaa	acacaaaaag	atgatttttt	6240
tttaaaacaa	gatttttaaaa	aaagcatgtg	ctacgcttcc	ttctgccact	aagcctacac	6300
atggctcctct	gactgaatth	ttcccctcat	tctgcttcat	ctaataatgtg	cttttcaaac	6360
ctggaattga	accagggact	tattcatgct	aggcaaatgc	tctaccatag	agctataccc	6420
ctccaactcc	catctcaaat	atcatttcca	aagacatttt	cttggtctct	tatttagatc	6480
aggtttcttt	gtcctcctgc	agctatgact	tcattccttc	agaacactcg	tcttagcttt	6540
aagttctgta	tttaattagt	attgttttca	ttctctctgc	tagaatgcac	tttcaataaa	6600
ggcaggtagc	cagccacagt	gcttaattaa	gcaacagccc	aacgatgtca	ttcactacat	6660
actgggacaa	gatgcctaac	atcatctgca	gataaagacg	aactactggt	gtcaggagac	6720
agctaagggg	tccagggcct	gggcacgctg	agtgtgagca	ctggagtcog	ggtgccaga	6780
aacgcacata	aatgcaatat	ggaatgggca	atctacctct	aattccttct	ttaagacagt	6840
ggctctccag	agcaagctgg	ctagcaagac	aagccatata	agtgaactct	gggcttgacc	6900
aagaccctgc	ctccaggtgt	aactcccaag	caaaaggatg	atggctcaca	aatctcaggc	6960
tatcatgttc	atgtacaaaa	tgtcaaccgg	catacacaca	tgacacacac	tgaaaactgg	7020
gagaaaaata	gaagaattgc	aaccaaaaaa	tgttaatttg	ggacacataa	ttgcaggcgg	7080
ggagtggggg	gatgacagaa	ggtgaactga	gtggaccgag	ggaaagctgt	gctagcggca	7140
atgagaagaa	gggtggggca	gtctgagcaa	gggttcagca	atcaccacgc	tttactgtct	7200
gcacagcctg	gctgtagaat	gctgggcttt	atcacacaga	attattcagt	atgtgtctatc	7260
tttacagtaa	agttattcta	tcaggctatg	ctacttcaat	agaacaagcc	tgaaaaagtg	7320
gtctgctgct	gagaacctga	caaagatgac	ctgttagaac	tgtctgccaa	gtgtggaatt	7380
ccagcactgg	ggaccaggag	ctcgagggtc	acccagatg	caggagtgta	gaggccagtc	7440
ttggcaacat	aacatcatgc	ttcagaaatt	aaaaacaaaa	nnnnnnnnnn	nnnnnnnnnn	7500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	7560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	7620
taaagtagat	agttgcaagc	atgaagaaag	acaaaccact	tgagcttttc	ttttgtcgta	7680
aggaggaaac	cagacaggct	cagagagatg	gctcagagat	taagagcact	gactgtctct	7740
ccgaaggtcc	tgagttcaaa	tcccagtaac	caatggtgg	ctcacaacca	tctgtacagc	7800
tacggtgtac	tacataacata	taataaataa	taataaataa	taataaataa	aatcttaaaa	7860
gaaaaaaaaa	aaaaaccta	ccaatcagcc	aggcgatggt	gacacatgtc	tttaatccca	7920
gcaacttggga	ggcagagaca	ggtggatttc	tgagttcgag	gccagcctgt	tcttcagagt	7980
gagttccagg	acagccaggg	tgatacacag	aaaccctgtc	tcaaaaaaca	aacaaacaaa	8040
caaaacaaaca	aacaaaaaag	gaggaagcca	gacaggatgc	actttatacg	tgaatggaat	8100
tgacaaaaga	caagtgttat	aagtgttagg	gaaaggggga	ggacaacggg	ggttcagtgc	8160
tgtggtggaa	cacgtattag	aaggctctgg	gtatcctggt	tccgacaaac	aggcactccc	8220
aatcacacag	gccactggat	gtctcaggca	gagaaagatg	tgatagattg	actttttaac	8280
aatcacagac	tgtgtggaaa	atatattgtaa	ggttgtcatt	gtcaccacag	atagagctga	8340
tggttattca	aacgaggatg	ggacaacaga	aatgggagag	agggatgtga	gaaccatttt	8400
gaaccagggt	gatttactgc	gcacgtgtat	agggtctaca	gggagtggga	tatgtagagg	8460
aggcctatgt	tcctaacttt	ggtaatgagc	ttattacagt	tactatgcac	agcctggaag	8520
atactggaaa	agggtcaggc	taggctagaa	aggtaactac	tgagggtttg	acagcccctt	8580
ggatgtcagg	atgcagcaag	cctacctctg	tatgtagtca	atcccctctc	aggctatggg	8640
tcctgcagat	catccgtctc	tgtatccatt	attcccagtc	catcctctga	gtggctccct	8700
cttatccagt	ttaacaaaa	gctgactgca	agctcccaag	cccagggtc	tggtcctttt	8760
actccttggt	attgtacttt	accctgtttg	cttggtgatg	agtgtgccct	ttataaacat	8820
ttgtgaaagg	gggaatgaag	aagaataann	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	8880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	8940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9000
aattcccagc	atcacatggg	cagctcacia	ctgtctgcaa	ttccagttcc	aggggattca	9060
acactcagaa	acataagtgt	aggcaatcta	cgtaacataa	aaataaataa	atgagctgga	9120
aaagaaaaca	tgtttcaaaa	tatacaagta	atggggctgg	aggagatgtc	tcaatgggta	9180
agatcatttg	ctgctctttt	ggaggttctg	ggttcaattc	ccaccaccca	catgacagct	9240
cacaactgtc	tgttaactttg	gtcctgtggg	agctgatgcc	ctcttctggt	gtgcagacat	9300
acatgtagac	aaaacacctg	catacataaa	ataagttttt	aaaaaagtta	cacatacacc	9360
cgtgtgtaat	ataacacaca	ctggcttaac	ttcctcagca	ctgactgttc	accatacggg	9420
ttcccactgag	gttttggttg	cattctatca	ccgaaaaaaa	aaaaaaataa	ttagaagaaa	9480
gtatatacat	ataaacctct	ccctaaaata	aagttttctt	ttctaaaagt	acatccttat	9540
ttttttatth	tttttttttt	ttagaagatg	ggaacaacag	ttctgtctac	actgtatttc	9600
tagcatgtaa	catcttgcaa	gtacttaacc	gtattctata	tcagctcaac	acacttacta	9660
ccgaagactc	aagatcacia	aaaaaaaaaa	aggacccaga	ctggataatt	aaacgtttct	9720
tttgtgttag	taagcgacct	cttccttaga	agatactaca	gtaatgtctga	agaaatgaca	9780
catctactgt	aatctgttct	ctgggattcc	aactgttttc	ctctgctact	cctcccttgg	9840
cggcaatggt	cgtctgcata	cggctgagct	cctcgctgcc	ttgttaaacc	tccttctctga	9900
acttccgacc	tgtagttccc	gctctacagt	gcaagcgagt	ggataaggaa	gcgcatacct	9960

gccgtctttc	aggggtgttg	cgatgaactt	gtggacctgg	cagacacagt	tgctggccag	10020
ctgccctccc	tcgaccaggg	tggtcagctg	cgtggccagc	atgaacgctg	caaaagcaga	10080
gagagagggg	ctcagtcctc	aagcctttcc	ttaaccggaa	agctcatcac	aaggagaaac	10140
attaaataca	gctgtttaaa	actcctccgc	cctgcagaga	ggaaagcagc	atcaatccgc	10200
cccatgtaaa	agtctgaggg	tcttcctaaa	tggtatctgt	ttctcacagt	ctccaaatca	10260
tttttactgt	aattctagtt	tctggggaaa	gacctttctc	ggtccttagc	cccgtgacta	10320
gagacaacag	gcaaatatct	cagaaaggcc	cccattttct	ttttaaaagt	tctannnnnn	10380
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	10440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	10500
ttcgggtccct	cgaatttggg	tttcttctgg	gacgtggtag	catgtgactg	tcactccagt	10560
gcttgagaca	gcagaggggt	caggaactcc	aggctggcat	tagctgcaga	gctggagcag	10620
gtcctggaga	acagaaactt	tggttcagc	attaatgaac	tagaagaatt	tttttgtctt	10680
ctgttaataa	taaataacct	cattatcttc	tcataaacag	tggtgccttt	ttattttaagt	10740
ttttaaggat	caggcacaga	gactccatgc	cagactacca	ctcaaccact	gagctacacc	10800
cccaacttgc	cttctctgta	ttttttaaat	tgtatcagtg	gccaccaaac	atggggagag	10860
gtcagggggc	tacgtggagg	aattgtttct	ctcctaccaa	gtgggcccc	ggtttcaaat	10920
tcaggtgacc	tggtctggca	gcaagcacct	ttacccttaa	gccatctcat	tggtctctac	10980
ttttaatggc	cccttccctc	gctctgaggg	aggctctccc	tatatagccc	tggtctggcct	11040
caggctcgca	ggtccaccag	tgagcaccag	gtttctgctt	gtccttacct	ccccagcact	11100
gtggttataa	gcatgtgcca	ctgtgtcaaa	ctcagtcact	aagctttgcc	aagccatagc	11160
ccagcccttg	agtttactgt	ttgtctgtgt	ggtgatttgt	caaaccactt	ttgttccact	11220
gaggtatttt	gtcaagtttg	acaaaattag	tttagtatgt	aggtcttttt	ttctggaatc	11280
ttctgttata	gtctagtctg	gtcctgaact	catgatcttg	cctcaacctc	acgattattg	11340
aggattattg	agatggacag	gctgtgtgac	catgctcggc	tgtgtgtttt	agcatgcatt	11400
agtcatttga	aaaacgttgg	ctcatgacac	tttacaggtc	ttccatgttt	gatatgtttt	11460
atttaatcca	aagtaattcc	agcaccagag	gctgagacag	gaggatctca	aggtcaacct	11520
agagatgcat	agcaggcggg	gccccactcg	gttaggttaa	tatcatcact	gacttcagga	11580
gaaaagtctt	aagtattggg	gactaaaagc	aggaggatct	gaagttcaag	gtcatcttta	11640
ggaaacttagc	agacttgagg	ccagcttggg	cgctgtggga	ccctgttttt	aaaccagaaa	11700
acaaattgaa	aggaaaaaaa	aaaaaagctg	gaggaaagtga	atgtgagtgt	tcacatagtc	11760
ctgtttccac	aagaaaacag	ggttactttt	ggcaacaaat	aggtgctttc	tttgaaggct	11820
ggcatttttg	tgacttgtca	ttggagaaat	gatttaatta	agacttttct	actgagtgcc	11880
tctgaagagg	ctcttttaaa	tttagtttaa	ttttatctca	ttgttagtgt	ggtgtgcttg	11940
tgcacacaga	aggcagcttt	ctagagtctt	ttcactctct	cctccacagc	tcctggagtc	12000
aaactcaggc	ctgggctagg	caagctctta	ggacagtgtt	agctgtagct	tattaagttt	12060
ttaagaattt	ttataagact	ctgtttttct	ttctcaggtc	atgatacagc	aggaaaatac	12120
atccataaag	cccacctctg	aggtcattgt	aagtaccggc	atgtgtgttt	agcataatga	12180
agatggttca	cttatagtta	attaaacatt	ggattggatg	gaagacatgt	agttttgggt	12240
acttcccaga	aacacaaatg	cacattcttn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12420
cttctattcc	ttcattacat	tactgttaat	agtggtaact	atgtaccaaa	gagtc aaata	12480
actcttgagc	catccaaggc	agaagggaag	ctggcaaaaa	tgtatgatga	tctgggatgg	12540
gaatgtactt	cagtttgtac	aggaggccct	tggttcattc	catttctggc	aatgcataga	12600
cctgtaggat	ctcagcactg	gtgggggggtg	gggggtgagg	gtgaaggggc	gggaggttaa	12660
aggcagaata	gtcataaatt	caaagtctgg	gtcctggaaa	gaggactaaa	cgattaagag	12720
ctttagctgt	tcttctagag	aacctggtgt	gatccccagc	acatggtgcc	tcacgactgt	12780
ccgaaactct	gattctaggg	ggatctgaaa	accctcttct	gccctctgta	gatacagaac	12840
acacatgggtg	cacatacata	catgcaaccc	aaacaaccca	tatacataaa	atattttttt	12900
ttcaaaaaga	cattcaaatt	cttctctcgg	tatatagtgt	ttaccaaacc	tcaaaaacaa	12960
aacaaaacaa	aacaaaacaa	agaatcatta	atgttttgcc	ttcatgtatg	tctgcccacc	13020
acggacatgc	ctggtaccca	gggagattaa	aagaagacat	tagctcccct	ggaatggaga	13080
taggtatgat	ctaccacttg	ggtgctggga	acctgggtcc	cctgcaaaag	cagtaaatct	13140
ttttaacccc	taagctgtct	ctcccaacgc	ctaaagattc	ttgtaacaca	gcatgatgag	13200
cactggcaag	catagcatgg	taatctgact	tcagggcggc	agattttgag	cttaattgctt	13260
gattattaga	agtaacgtac	tagatttaat	gcctggagct	tcaagcaaca	aaattaactg	13320
aagaataaaa	ataaaaaacc	tgccagccat	gatgttaatc	ccagaacttg	agaggcagag	13380
gcagggtgatc	tctgtgtttt	gcaaggccag	ccacaatcta	catagcacgt	tgcatgannn	13440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	13500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	13560
agaaattcaa	agaaacaaat	gccaaataaa	tacacataat	gtaataaaga	gataattgtc	13620
taaaaaactc	aaggctttta	atggtaagat	atcatattct	tggtatgaaa	gatctaattg	13680
caaaatatat	caatttaatg	caattatgta	tattcaggag	atctctggtt	ggcttttgaa	13740
cttgatagca	ctcttataat	tcacatagaa	gaaaaaaaac	catgaaaact	gccaaacatt	13800

attagaatac	tccacagatg	gtatttttggc	agcacataca	tcgaagggtc	gtgaaagatg	13860
tgtagatcat	ccacgccttg	ctagggagag	ggcgggtgtg	tgtggggggg	atagctgttt	13920
gggaaaataa	cctggtaatt	cctcattagt	taaatcatag	tcagaacctg	gactagcaac	13980
ttctctctaa	aatacattca	ccctcagcat	ctgcattgcc	aggaaaccac	tcctagcagg	14040
atctgtacgt	ggatcaaggt	agtagcatct	gcattttaatt	gacattctcc	taaatgtctt	14100
aaatttatctc	tagattactt	atagtagcca	agatgatgca	aattatgtta	cactgtatta	14160
tctggggcgt	aacaagaaaa	tgtctctact	caggttcatt	cagggtgcagt	acttcccctg	14220
aatactttctg	aatacacgga	tcaagaagcc	acagaaagag	ggctaaccat	atacaagcat	14280
atagtacact	aataaccatg	tacaaccata	tagtacacta	atattcagtg	cattactcaa	14340
aatgcaaaca	gatggaaaca	atccaacagc	ctgtaagctg	aaaaacaaga	taagcaaaat	14400
gtgctggggc	tagaggccca	ggtctataat	tccaactaag	gtcgaggcag	gaggtagctca	14460
agttcaaggc	cagcctagac	aacttagcaa	gacctgtgtc	caaaacaaaa	agtaaagagg	14520
ctgaggatat	agctcagtat	agagcatctg	cttagcatgt	gcactgacag	ccgtatcaca	14580
gaggaaaaaa	aaaaataagc	aaaatgtgat	ctgtctgcac	aacaggatat	cacagccccc	14640
taccacaggg	gaacgacaca	gtaacacaac	aaaaacttag	ccctgaaaat	actatggtaa	14700
ataaagaagt	gtcactgagg	atcaggaaat	gcatgactcc	atttacatta	tatagaaatg	14760
agaagatcag	tgagcctcta	ggactcaaga	gatttgggat	tggcagctaa	agggtagctg	14820
gtttctttat	gggggtaaga	aaacatttcta	aacttaactg	tgagaatgac	tactcaacaa	14880
tgtcaagtgt	tcaaaaatca	tacttttttt	tttttttggg	ttttcaagac	agggtttctc	14940
tgtgcagtc	tggaaactcac	tctgtagacc	aggctggcct	cgaattcaga	gattcacctg	15000
cctctgcctc	ccaagtgtctg	ggattacagg	catgcgccac	cattgtccgg	ctcaaaatca	15060
tactttttaa	aattgcccag	tgactcatga	atacaatcag	aggcgggaga	ggacagtggc	15120
aaactcagga	taccagtgtc	ttttatgtct	gctgcccaac	tatcaatttc	ccatagttac	15180
cagagaactt	tttggtttgt	ttcatcttat	ttgttgcttt	tggtagaatc	tcaatatagt	15240
aagatacaag	gctggcctca	tactatatag	ctgaggacga	ctttgaactt	ctaactctcc	15300
tgcttccatc	tcccaagtgg	tgggattaca	ggggtgtacc	gctatgcccc	gcaagcacia	15360
agccatttga	accacacccc	agccttttca	gagaaacctg	tacaagcctt	agtgccttag	15420
catattaagg	caacaaaaga	cataatgcgt	ggctaccata	gagtgtttgc	ctaccatgtg	15480
tgaggctcta	gctctaaatg	ccagcactta	taaaaaagag	ttaaaaacac	tcattgactca	15540
aggatgacta	tgcagtcttg	tgtacaaagc	cccgcattca	atccccagca	ccgtgcacat	15600
caggcaggct	ctgtagagga	cccagcttaa	ggtcatcctt	aggtaagtta	gaggccttag	15660
atggctacat	tagatgagac	cctttctcat	aaacagaata	aataatttta	agctcctgat	15720
caaacactat	gccttcccat	cacactcaga	ataaagcact	ctactggccc	tttaaggact	15780
gcccatctgg	aagagaaacc	taagtacatg	tccttgcttg	tgtcatatgt	gataacaaac	15840
tcactggaaa	tacgaaaata	cagtcttaag	cttggtcaga	aagcttcccc	agcaacatga	15900
tntcagagga	cataatgcag	aaagtggaca	aatgcaaann	nnnnnnnnnn	nnnnnnnnnn	15960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	16020
nnnnnnnnnn	nnnnnnnnna	tcagaggaca	tctttcagga	gtagttctt	tcctccctct	16080
atagctttca	gggatcaaac	tcaagtgtgt	actgagcgct	tatgcccagt	gcgccatcgc	16140
accaggctcg	cttctttggt	ttttatgggt	ctgaatcaat	tagcaccatt	acaacaatgt	16200
tgacaatcag	caagtacctt	tctctacctg	gctagtaaga	gaagtaagt	cctttgggtg	16260
gtgaacgcag	tttctcttgt	gaagtgcagt	gacttgatct	ttgctcacia	cgttttttag	16320
gtccttaagt	tgcttgggtt	ttatgggaaa	ggctcttggg	ttttttgaaa	agattttact	16380
acaacttgat	ataatcatta	tttttaatcc	tttaaatagt	atgacttatt	ttaacagatt	16440
aatattgaac	tgttctttca	ttctttacata	ataaatcctg	ccttaaaaaa	taactctctt	16500
agcttccttt	ctctattttc	aaatttggtt	tatatttttg	catgattttg	aacattttata	16560
aaagtaggca	gacaacacag	tagaaccaag	tcccatata	gctgtgcaca	tagcttcaga	16620
ttattgcctg	ataataggct	ctgtttttgt	tctgttttct	cacagggatg	tgttattgtg	16680
tgtgtgtaca	catacatata	tatatgtatg	tatgtatgta	tgtatgtatg	taatgaactc	16740
cctttaacaa	aacaagtact	gggctgggaa	gacagcacag	ttagttagtg	gtttaaccgc	16800
acaagcatga	caaccagagt	tgagatcccc	accaaccgca	taaaaagctg	ggcatagtgg	16860
cattgacctg	tagccctggg	gctggatgaa	agctggggag	gcaggtagat	cggcagagct	16920
tactggcaac	aaatctgccc	agtaggtaag	ctctgggctc	agacatccta	tataggaaaa	16980
agatgaagg	cgaggcgag	cggcacacac	ctttcgtggg	agtgttgag	aggcaggggc	17040
aggccagtct	ctgtgaccag	cagcctggcc	tacatgtcaa	gttgacaggc	agccagagcc	17100
accacttact	gagactgtct	cagaaataag	ttttttaaaa	aattgagatg	aaggagctgg	17160
taagatggct	tagaaggtaa	aggcacttat	cactaagcct	gaagccccga	gtttgacctt	17220
ggacccccaca	ctgtagaacc	aactcctcca	agttcttctc	agacctccag	cagagcacia	17280
gtgtatgcag	acacacacac	taagtaagtg	aatgtaaaaa	acatgacgta	gtggcactgg	17340
ccttttaaac	cagcattggg	aggaagaagc	gggtggatct	cttgagtgtg	agaccagctt	17400
agcctacata	aggaatttca	aggcagccag	ggctaccctag	aaagttagctg	tttatgaatg	17460
aatgaataga	accgaaggag	aaagagagac	agacttaaaa	aatatatgct	ggagagtaac	17520
agaagaggac	accggcttgc	tgggtgtctg	acctctggct	tgtacacata	cacatgtgta	17580
gtgcatacac	ccacatacaa	ttgtactcag	acacacacaa	acatgtactc	attcatatac	17640

tgacacacctc	aacactcaga	aaatgaaaaa	acaggtacca	tttacacctc	cgtgttcggt	17700
ttccaaccac	tcatatgtat	gggttgtaaa	tgcttatatc	tgatgtgtc	tgatatttg	17760
tgataacatt	caaagttgag	tcaggatcca	acgtaaacctt	ggatagtagt	gggttgatgg	17820
tctggaagcc	tgctcgcagc	tgcttttttc	tcctcgtacc	ttttccctcg	ttgttttcta	17880
cgacagcagg	tcatttgtct	ctaagtgtta	gtttcccatc	ctctctcttt	tgctgtaggt	17940
agccttgtag	tagtcacctg	tgctctctgt	aaaatggctt	tgccgtgtta	tttcaatatg	18000
ctatcatcct	catcttgcta	tatttcattc	aatatatgta	tatatataca	gatagattaa	18060
aattattttta	attttatgct	tatgaatggt	ttgcctaagt	atattgcacc	ttgtgtgtct	18120
agtgtccaca	gaactcagaa	gaaagtgtca	catattctgg	aactggaatt	gcaggtggtt	18180
gtaagccacc	atgtgggacc	tggaaccaa	atccaggcgc	cannnnnnnn	nnnnnnnnnn	18240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	18300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	18360
cagacttctc	gtttctctca	tggaacaggt	tccccctgag	atttactagt	ggaagaaagg	18420
cactcaaaaa	gcagggagcc	ctcgtacaaa	tgagacttcc	tagctatata	attaggccga	18480
gatgcacaca	tccacagtca	ttacccttct	tcagagcctt	tgctatgtca	agtgtatttc	18540
gcccattgtga	acttttagaac	tggtctgttg	tgtttctata	aaagtgtagt	tggtctcttg	18600
attgggattg	tggttaaattt	atagattcat	ttggggagag	ctgacaactt	tacagtatta	18660
aatgtttttca	tccgcggaaa	aggttgtctt	gccacttact	tggtcttctt	tttatgccct	18720
taagtaaagg	tttatagttt	tctttatata	agtcttgcac	atttctctgt	agatttatatt	18780
ttgcttgtaa	ggtttctctg	gttggtattg	tgtataaaca	ttttctccca	attacatacc	18840
ataattgatg	gttaggagta	taaataaaag	gtagaatttt	aaaattgctg	tatgaatgac	18900
tgcttttgcta	taattcaaca	tcttttcatt	tttatgcca	tctacttgac	atgttttaggt	18960
taaatgatga	tatctgtaca	gagtaatcct	ctgatccagt	atttgcacat	cttactttct	19020
aacgtccata	gcatagatac	acatcttata	ctgttgagta	catatatatt	taaggatttt	19080
accatagtct	tataatatgc	agcgtgcttt	ggttcaagac	agtggccctg	tgttcctcaa	19140
cattaacatt	tttttcatca	caaatacaca	ttgaccttta	tcaaatTTTT	aaaactatct	19200
tgagagaaat	gaccattttt	cttaatctgt	taatgtaaaa	tttttataaa	aatagttata	19260
aataatatta	gcttacatat	ttcttctggt	ctttttttca	actcttagaa	tcagagtatg	19320
gtagtctcag	actaaaccag	gagcttccta	tctgtttctc	tggtcttaag	tcacttatat	19380
aatgtaagg	tgctgtgtat	atctgccagc	taggccttat	atacaaaagg	caccatcac	19440
aaacttctaa	aacagtctta	ccacttagag	accatgttca	aacatatggg	cctttgaggt	19500
aattgccaca	ttcaagctat	aatattgtta	tctaaggga	tatcttcaact	tctagcagat	19560
gcctaaaaat	atctaaagg	aaacactggt	aattgctgtg	ttgtttgatg	ctgctcttcc	19620
tcctctcctc	ctcctcttct	ctcctcttct	ctcctctcct	ctcctctcct	ctcctgcttc	19680
tccttttctt	catcctcctt	tcttttctta	tttttgaggc	atgatttcac	catgtagccc	19740
taggtaaccc	gtaacttact	atgtatgtag	accaggctag	cctctgtctc	ctgagtgtct	19800
atattaaagg	tggtgtatcac	catatccagc	aacacttgct	ttgagatggt	tagaggaaaa	19860
aaaaatatac	gtaaataaag	atggatgcca	attactaaat	tggtacttcc	agtcacactt	19920
tgtagctagt	tcaaggccaa	aatagggtat	tggtttctac	tttgcaagtt	ggctccatta	19980
agaggctttt	cttctcttgg	tctcactaga	taggaaggag	agagaggagg	gaaggagaga	20040
aagcggttga	ggagtgggag	gtagtgtgac	cgagaatacc	cagtaggctc	atatatttaa	20100
atatttggtc	cctagttgat	agaactgttt	agaagatta	ggaagcatgt	ccttaggggt	20160
ttgaggtttc	aaaatttaat	gctagaccca	gtctttcaag	ggagggggag	gtctgtctct	20220
ctctgcctgc	tgcatgcaga	gctctcagct	actactctag	tgtaagcct	gtgtgcttcc	20280
tgactcaatg	acataaaatt	aagctgtaagc	aagcctccaa	ttaaatgctt	tcttttatag	20340
ttaccgtgat	catggtgtct	cttcacagaa	atagtaacct	gtggtgattt	taatatgcct	20400
ggaccaggga	gtggcacttt	taggaggaat	ggccttggtta	agagggaagt	tgtctctgtg	20460
ggggtgggca	atgagaccct	cgtcctaacc	atgtgagaac	cactcttctc	ctattggcct	20520
tcagatgaag	atgtagaact	ctcagatcca	cctgcaccat	gtctgcctgg	aagctgcctt	20580
tgttcccacc	ttgtgcctcc	aattaaatgt	gttacttata	agaattgttt	ttgggggggt	20640
ggagagatgg	ctcagcagtt	aagagtactg	actgctcttc	cagaggtcct	gagttcaatt	20700
cccagcaacc	acatggtggc	tcacaacctt	ctgtaatggg	atctgatgcc	accttctggt	20760
gtgtcagaag	acaggacagt	ataccacat	acattaaata	aataaataaa	taaaataata	20820
aattcttttt	aaaaaagaat	tgctttgggt	atgggtgtctg	ttcacagcag	taaaacccta	20880
acataaccct	gactaagaca	acaagtggag	aaagggtgtg	tgtagacttc	tggtctctctg	20940
gaagctcacc	tcnagcatgaa	gcttgcgaa	cggnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	21000
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	21060
nnnnnnnnnn	nnnatcctaag	tactactgtac	tgcttctcaga	cacaccagaa	gagggtgtca	21120
gatctcatga	cagaggttgt	gaactcagac	ctttggaaga	gcaatcagtg	ctcttaactg	21180
ctgagcatct	ctccagccca	aaataattct	tactagtaac	atggaacaat	caagttttat	21240
tatatgatatac	atattaatca	acttataagt	actatgattat	gcacatttat	catatcgtgc	21300
aaccatcact	gctgtcgttt	tggtttgttt	tggtcttttg	aggcccggtt	tctgtgtgtg	21360
tctggaactc	actctgtaga	ccaggctggt	cttgaactca	atgatctgcc	tgctctgccc	21420
tccaagtgc	tgaaaacaaa	tgtgtgcacc	accacctctg	gctatcactg	ctgtcttttt	21480

ttttttttta	acagttatatt	atttcgtgca	tgcattgtgtg	tataagcatg	taacgtatgc	21540
catggtatgc	atgtggagggt	cagaggacaa	ctttcaggag	ttagttcttt	cctcccactg	21600
tgggttctag	gaaccaagct	caggttggtta	gacttgcatg	gcaagtgcct	ttaccacaga	21660
gccatcctgc	tggccctact	ataggtcctt	atataaaaaa	atcatatgcc	gggcaaaaac	21720
caaaacaaaa	ataaacctca	aaaaacaaaa	ggaccatata	atattgtggg	ggagtggatg	21780
aagtcctgaa	cgaatgtggt	ctggtgacat	gtctgtactt	cagacccatg	ggaattggca	21840
aagccttcc	ctgggtcctgt	gaggatgctg	atagtctgtc	taaaaactag	agatcacagc	21900
tttctcctct	ggatgactgt	aaccccgat	tgttcctctt	cagagactgt	ccaccaagct	21960
accctgccta	cttaagctgt	acacaatgaa	tgagctgagt	ttccagggtta	cagcacagta	22020
gacactgtcc	atcagtgaga	gcacagccta	gcctaacagt	acacatgtct	gctttcttca	22080
cgtttccaga	accaagcctt	gctggataga	gcataattgt	ctgtttggtc	tatttcactt	22140
gataaaaaag	tttcaaggag	ggccagggtg	ggtggcacac	gcctttagtc	ccagcactcg	22200
ggaggcagag	gcaggcaaat	ttctgagttc	gatgccagcc	tggtctacaa	agtgagttcc	22260
aggacagcca	gggtatata	gagaaacct	gtctcaaaaa	acaaaaaaaa	acaaaaacaa	22320
aacaaacaaa	caaaacaaaca	aaaagccaaa	aatccaaacc	cccccaaaaa	aaaaacaaaa	22380
ccaaaaacca	aaaaacaaca	acaacaaaaa	gtttttgagg	tttaatttat	tgcatgtcac	22440
agaattttc	tgtttaaaaa	aatggctgaa	taataattca	ctatccattc	acgtatttgt	22500
aggcattcat	gtgtgtagt	gttttaataa	aaatagcccc	cataggcttc	tacagttgaa	22560
tgcttagtca	ttgagtagca	gtactagaga	gggaattgaa	ggtgtggcct	tattggagta	22620
ggagtggcct	tggtgcagga	attgtgtcac	tttgagggtcc	cagcaacaag	gttgctctga	22680
tcacatccaa	agacattcta	ggtctatgtg	atctggctgg	aattcagaca	tgcccttaat	22740
acacaccttt	aatcccaaac	aatgaaggta	aagttagttt	ataaaaagaa	gcacccatgt	22800
ttgaaaagtga	cgtttaatta	agagtgatga	attagagaaa	gatctgctgt	cacagagcag	22860
agaggaaaga	gaggcagcat	aagagggagc	atggcagagg	gagagggagg	aggggttttc	22920
accagggcat	ttgtacagag	acaggttgca	gagctagaac	aggtgaagac	agaacaagcc	22980
agagaatgag	aaggagccag	gagattagga	cagattgccca	atgttaatat	gctaagcaga	23040
gcatttttagt	cagaaaactga	gagaagtcaa	attgaatcag	ttagcttgga	aaggagtttg	23100
agcagcaaca	gctgagttaa	actagccaac	agaactcaga	aagaactaga	aaagatgagc	23160
ttactcagca	gcaaatctca	gaggctaaaa	acatcttaga	cctagattag	actgcatgga	23220
ggctagacgc	ttccagggtc	aggcctaggt	tagcagacgg	agagagtaat	aagccttgga	23280
gacaacagtt	aatacagaag	actatgtaca	gacatggata	tgaacctctc	agccacttct	23340
ccagcgtcat	gcctgtctgc	attgttagga	gtcatctagg	aaaggctaag	ggcaggcaag	23400
caacttttcc	agagatggtc	cactgttttt	tgcatggcct	ttgagaggcg	agctctgaga	23460
gggaagggtc	caagagactt	catcccagga	tgcttgctta	attacgacat	gccttttctt	23520
gtcactgtta	tttagtataa	tgactcctga	gctttagccc	atcctatttg	gcatatttcc	23580
tgcatagcaa	cataaagatg	aactttcaca	aattaatgct	gtttagatga	ataaatgatt	23640
ttataaaatt	cctgatttga	tttaaataat	tttaggaaga	aagcttttag	agatagttta	23700
gttggttttc	cagaaagatg	taataacgtc	agaatcaaga	atagaatgtg	gctgggcagt	23760
ggtggcagat	gccttttaac	ctagcacttc	ggaggcagag	ataggcggat	ttctgagttc	23820
gaggcacagc	tggtctacag	agtcagttcc	aggacagcca	gggctacaca	gagaaaacct	23880
gtcttgaaaa	acaaaaacaaa	aagaaaagta	agtaaaaggct	gcataataaa	gaatacaatg	23940
agctttcaca	actacaccaa	aaagagacat	gcttgggaca	aatttgtgat	caaggaaaaa	24000
tattcattct	agatcaggtc	caaggatgaa	gccacaagtg	tggtgatatga	tgaacaagac	24060
catggataaa	ctgttggttt	gagcttaaa	aataaaacac	tgctttgaaa	taaactatca	24120
acattctact	gtaactttcc	ttttataaaa	ttttatctat	gagataattt	tctaagaagc	24180
ttgtgtctat	aaaggtagat	aaggacagag	agaaagaaat	aagggtgtggc	atctgggctc	24240
tgctccatcc	acccaaataa	atatgtgtgt	gtgtgtatgt	atgtatgtat	gtttatctat	24300
atgtatgtat	atacatatcat	gtgtaggtag	gtatatgtgt	atgtatatata	gtatgcatga	24360
acacttgga	agttgatgag	acaagtgaga	ggttgggccc	ccnnnnnnnn	nnnnnnnnnn	24420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24480
nnnnnnnnnn	nnnnnnnnnn	nngaattcac	tctgtaaac	atgctggcct	tgaacctcaga	24540
gaaccgtgtg	cctctgcctc	taaagtgtctg	ggattaaagc	atgtaccacc	acaacccagc	24600
tagttttaa	gtttcttatt	tttttggttta	tgggtctttt	acctgtatgt	atgtgtgtgc	24660
accatgtgga	tgcatgggtc	ccttagagtc	cagaagaggg	tatcagatcc	cctggaaactg	24720
gagtgcagaga	gggttggtgag	ctgggacttg	aacctaggac	ttctaaaaga	gcagcagggtg	24780
ctcttaaat	ctgagcctta	tctccaggcc	gtcccatgga	tttggggggc	tttgtttcat	24840
tttattttgt	tttgagacag	ggtgtgtagc	tcattgcttga	atttactatg	aagccctgac	24900
tccctcaaaa	gtaaagatcc	tcttgctctc	gtctacagct	gctaggattc	gaggtcttgt	24960
accacatgct	cagcacagcc	atgattcata	acaataaaaa	aagaaagaga	gacctaaatg	25020
gccttagaga	taaaataaatt	attttttttt	taaagattta	tttattttatt	tattacatgt	25080
aatgcacact	gtagctgtct	tcagaccccc	cagaagaggg	agtcagatct	cattacagat	25140
ggtgtgtgag	caccatgtgg	ttgtctgggt	ttgaaacttcg	gaccttcgga	agagcagtcg	25200
ggtgtctcta	ccactgagc	catctcacca	gccccgagat	aaataaatta	taatgtatgc	25260
gtaagggtgg	atcatctcag	tctccgggaa	tcttgctgtg	tactccttcg	ctctcccttc	25320

tattcatgct	tgggtaactg	gccctggctg	attgatgaga	gctgatttcc	ccactgcctt	25380
gtggcaggga	ccactgcgcc	cacagggctc	cctcaggatc	ctcagtagag	agctgcacag	25440
ctgggtggaa	gtagagggct	gcatatataa	cacgatctca	actttatttc	tttaaataaa	25500
aattttat	aaattttata	cagctctata	taaacgaagg	aactattgaa	ggttcagcaa	25560
ggacctgcca	acggttgta	agggtaattg	cgatgtagt	atTTTTTTT	cccttccat	25620
tttacttcca	tactttctac	attacccac	aactggcaag	tattatttta	aaatgaaagt	25680
aaatagtac	agatgacttt	gaaggaaaat	tgaatcggt	aaaagaaagc	tgagagacca	25740
cccgggaagc	ccaggctaaa	tgtaatctgg	gtcaggcctc	ccaggcctgg	ggtctcaaga	25800
tggtcagctg	agggaccctg	gtgaccctct	tgggccagca	gggacgggga	ggagccggaa	25860
gctgagtacc	caaagtgtc	ctctgggctc	caagggcctg	cacagagact	gtgtgggaat	25920
caaaggatac	aggctagagg	actgaggcct	gacgaaccca	gctatcattc	gtcctagaac	25980
aggaggcaga	gctccaagag	tccaacaaag	aggcaggaag	ctttgggacc	cgagatgggc	26040
gatgggatta	gaaaggcatg	tttgcaaata	ctttcaaatt	tacgatgcac	actcactgga	26100
aacccacccc	ctgggtgtcc	cttccctgcc	tcttgccaca	cccaatagct	gacatcactg	26160
gagaaagtcc	caagaccagg	ctggctggag	ctcctgatag	gttccaccct	cctgcagagg	26220
gccctcgaag	actagcttgc	tcgcccacac	cgccagatgt	ctgtgtcttt	ctctcttttg	26280
ctcccaaccc	tcgtctcttc	ctccaacctc	agtgagggtt	cccctgcttc	ctggggaaag	26340
tagaacttgc	cagtgtctac	tgtaatgtcg	tccctgtagg	tgatcatggtc	ccccattact	26400
gggagcaggt	atgcctcaga	tctccctcta	tctcgtgccc	tttcaggctg	tctcagtttc	26460
tctctgacag	ttcctctcct	cctgaatcct	gcttggtggc	atgcgaacag	gctcaatatc	26520
ttccatctca	aaaaacaaac	actgggaagg	gtttgagaga	cagagagcat	gggtaattggg	26580
tgccccagct	tggctgggaa	ggggtaactt	acaatgctct	actgcccagt	agggtagctg	26640
cagttgtcaa	tttaattgtaa	atttcaaaat	agctagtaga	gaggatttta	gatgttccca	26700
atcccaacac	aaagaaatga	taaacattca	aggcgatggg	tatgctaatt	gctctgatct	26760
gatcacccga	cattgtatac	atgtttttga	aatgtcaggc	tgtaccccat	aaatatgtac	26820
aattaccgtg	cagtgtattca	agataaaaac	tataatttta	aaaagctaaa	aacagaagga	26880
aatagctgcc	cttgaccccc	ccacccccac	aaggctcttc	ctgtttgtcc	agccacttaa	26940
tgtcagagct	tcctgtggga	gggtggtttt	ggtgtacaca	gacactcctt	cctccctcct	27000
tccccataag	aggagtcacc	cctgtcccac	gatgccatgc	agggccacat	gcgtgatatt	27060
aaccagtaag	atgtgagcag	ggatgatacc	tgtctcttat	aacaaacgga	aaaaaaacca	27120
caccaaacca	aaaacaaaca	aacaaacaaa	caaacaaaaa	cagggttggt	ctgtccctgt	27180
gtcttttccc	acataaagtt	aagcacacaa	agtagccacc	atTTTtttat	ttgtcccttc	27240
ccccaccctt	ccccgagaca	atgtttctct	gtataacagc	cctagctgtc	ttggaactca	27300
ttttgtagac	caggctggcc	tggaactcac	agagacacag	agattcacct	gcctctgcct	27360
cccaatgca	gggattaaaa	gcatgagcca	cgaactaacc	agtacccag	agctcttgac	27420
tctagctgca	tacgtatcaa	aagatgacct	agttggccat	cactggaaag	agaggcccat	27480
tggaacgcga	aactgtatat	gcctcagtac	aggggaacgc	cagggccaaa	aaaatgggaa	27540
tgggtgggta	gggaagtggg	ggggagggta	ttgnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27660
nnnnnnnnnn	nnnggcagga	tcctgtgttc	atgtgcaaca	ctcgatgcaa	gctgtgtagt	27720
gtttgtttct	gagcacctga	aggggaccaa	gcaggctgat	gcccaggcca	cggttctctt	27780
ctggcccccac	tgcccaactcc	caccctctgg	catccccatg	atgaacatgg	ccacagatca	27840
cactactctg	ctcctctccc	agatccacgg	agccataggg	tcccagatt	catctctgca	27900
gctaacaagc	tgggcagtg	cacctccctc	aaggttccct	tcctgtctctg	agcagcagtg	27960
tctcccacag	tgagacactc	atgtccactg	gaagatatgt	tagccattaa	attcctgtgc	28020
taaaataact	agggggactt	gtcaatcact	acactcttag	ccccggactt	ctgactcata	28080
gaggggtggg	acagctcagg	gacctgcatt	ctaccaaata	gccatgtgtc	cctgatggag	28140
gaactgcccc	tggaacaacct	ctgcagcaac	tgaacctctc	gtggtctcct	agttcttctg	28200
gacaggtgtg	acccacgtac	ctagtgccag	gtgagagagt	gctagggcca	cactaagggg	28260
tgacaggaca	aggttgagc	tggtagatgt	ttgggcccac	aaagagaaca	ggtcagtagt	28320
aaaagccatc	atggcctgag	ccagcctcgc	agtcctctct	gcagttggga	cactcttgca	28380
gtgtcctggg	gacctcttga	gggtagcatg	gtcaccaaaa	tcctacaagg	acagatcaga	28440
agtcagttag	gtcaagggaa	cagctctagg	ttctctgtgt	ccctcacgga	cctttttttt	28500
tttttttttt	tttttaagat	ttattttatt	attatatgta	agtacactgt	agctgtcttc	28560
agacagctcc	agaagagggc	atcagatttc	gttagcagtg	gttgtgagcc	accatgtggg	28620
tgtctggatt	tgaactcagg	accttcggaa	gagcagtcgg	tgctcttaac	cactgagcca	28680
tctctccagc	cccctctcag	tcctgatcgc	acagggcagc	aaaggccttg	tcccagatct	28740
gaggagagtc	atgctgaagt	ccttccctacc	ccaccccttc	cgaacccctg	aacatcagcc	28800
ccataactac	tgactcccc	acccccattc	ccttgcttcc	actgatccgg	tcctcctctt	28860
ccctctggcc	ccacccattc	ttccccagcc	ccacctgatt	gtacctgggt	gtccaaacttg	28920
aaaggggcag	gcaggggcag	cttctgtctg	gctctgtcac	tcactggctg	tagaaaatgag	28980
aaaggagatg	aagaaaaagg	ccttccccatg	ggctcccatc	ttgccaagac	ataggtgagt	29040
ccctttggct	cttcccccta	aacctctcac	ttttgagtac	ctgctggccc	gggagatcca	29100
cggcgctcac	cggagagaac	tggttgagaaa	agggagaaca	gagaactcag	cggttctccc	29160

tctccaccct	tctggcctct	cccagatttg	ccccgcctcc	ccagcatctc	cttcagcctg	29220
actgaccact	tcccactcag	acctcagctc	tgccctcaccg	tgaaacaggg	accttgacgg	29280
caggacaagc	tgagtacgag	gagcccccg	agcagtgcga	tggttcctgta	tccagaacag	29340
ggagtgttag	ttcctacctc	acgctcgaag	gccaagcagt	agactgctat	ccatgggttc	29400
cttgaccgga	ccaggctgcg	gaacctggac	tcaaaacata	gcagctgtgg	acctcactca	29460
ctctgagagg	tggtatttcc	ataagctttt	tttttcacct	gtacatttag	tcttcattct	29520
tttctgttta	caactgtggat	cagtctctgg	ttcaaattta	aagccctcat	cttgcaagag	29580
gaccttgccg	atctcccttc	atgccttttg	ctttaccctg	tcttggtaat	tcatggcaga	29640
agttcttcct	gctcccatgt	agatgttgag	gacccaaata	agaatctctg	taaatactga	29700
gcatgatgcc	tggtcccccac	cctagcaaag	ccacctgacc	tggtgttcat	ttcatccagc	29760
ctttctcagg	ctgccctggt	cctacccaaa	ggctctgaga	gctaactctgg	gctggcaggg	29820
cagccaagaa	cttctttgtt	gaccaatgaa	tgactggccc	agacaccttt	ggacttacgg	29880
gaactacaag	cctcatccca	cttctgtctc	aagtctctgat	ccagggtgct	tcggggaagc	29940
ccagctggcg	gaagggggga	ggctctcagc	ctagagagcc	ttcctttcca	tcctcagccc	30000
cctaccagg	ccttatttca	ggcaccagct	cttctaaaag	gtccttctgt	tatccctaga	30060
cctccacaac	tgtgttcaag	aaccttcagc	cagggcctca	tctccaatct	ggatatatga	30120
ttttctcgc	caagagttag	cctccaggtt	ttggagtctt	agagggttct	cctggagctg	30180
cctggacctc	tgctcctcac	cacccagga	cgctgtgaag	ctgcaggctc	cctgaataaa	30240
ttcatccaga	ccccttgcca	aggtgccagc	tgtctacttc	ctctgctgcc	caagcagcag	30300
gctgcaccac	ccctccatcc	tacctcttca	ggcttcttag	cgcagcacac	gcagcacacg	30360
gtgttctcct	ggaccagctt	gctcccccag	ctccccagct	gcagccagca	gggcttagct	30420
ctctcctccc	acaggacctt	tgctctcagc	caacccccgt	tcagcttggt	ttcagtgctg	30480
gtaaatattg	acctgtacat	ccggttaaac	attgatattg	gggccagaag	accctttccc	30540
atcaaggcta	cccagaaccc	tgcttgagcc	tgagagaagg	gtttacagga	gcagataagt	30600
gaggaggttg	ggcctggcaa	gccttctaata	gatccctcaa	cataggggat	tatccacagt	30660
cagtgaaggct	cagagaggct	gtgtggcctg	tgtaagggcg	cagagtgggc	tccagagtca	30720
cagccaaagt	cccaccacca	ccaccaccac	caccaccacc	accaccacca	ctaccaccac	30780
caccaccacc	accaccacca	ccaccaccac	caccaccact	accaccacca	ccaccaccac	30840
caccaccacc	accaccacca	ccaccaccac	cacctcatct	accatacta	anttgaggct	30900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	30960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31080
gaacatctct	ccttttttcc	tttttgagac	aatcttacta	tatttaggct	ccccttgaac	31140
gtgtgatcct	tctgctctg	cctcccaagt	gctgagggtta	caggcatgca	cagtcacatc	31200
tgctacaca	atgtcttagc	agcccttagc	agcaccaggg	gtcaggaagc	cctcaactgt	31260
cccttttagct	ggcttctctt	gtgaagggtc	atgtcttctt	ccccttccca	gcagtgagac	31320
ggctcttagc	cccagagcct	tccttccctc	agggttaaaca	gcaccagttt	ttgggtgggac	31380
ctcccatttc	cctctatctc	cctaagcaac	gaccttttct	gctctgactc	tcatctggga	31440
cttgaccac	aagacaaaac	tgacgcctgg	gctgtgtgtc	ctcgcacatc	attcctgtgc	31500
ccccctggag	tcaggtctag	gggaggaaga	cagggttcac	gactcagaaa	agaccactgg	31560
ctgtcctagt	gtgccctcac	ccatccctata	gcacgcacat	gctgatgtgc	cccctccgct	31620
ccatcaccat	cctctcatgt	acacgtgccc	tcctctcgca	gacacatgca	tcactaactt	31680
ttctgacttc	ccagaaaaat	atctgatctg	agaagttagg	agtctgccat	catcagctat	31740
ggctcctaaa	attaagtcag	acaatccatg	ggacatgaag	ggcaacaacg	agaagactcc	31800
tcgttctctg	ttcactctgc	ttttggcagc	accaccagca	ggaaccaacc	tggtctctcc	31860
taatccctca	tctatagcag	gtctcccggt	gggaatttta	gggacctctg	tggtctcatc	31920
caggggcaact	gccactcagc	tgctcagggg	gagacccctt	agaacaacaa	agaaatcaat	31980
gcagatttag	gcttcttggt	tccttcccca	gcccctccca	tcacaggcaa	cagcctccct	32040
tggtctgagc	tcaggaggct	gatttatcag	agagggtgct	agagaggcac	ctctgggtccc	32100
tctgggtagg	tagcaactga	gacaggagga	gatgggtcacc	ctgggcatcc	tctaccagga	32160
agtaaatgag	atacccttgc	agatgggacc	cctgaagttc	ctcccggggg	cgggggtggt	32220
gggtggtgga	gtctaagtca	cagatctttg	ttaccacgtg	gttagactga	ggactgaatc	32280
tgaggtggga	aatctgatgt	gcagggggaa	acacagaggt	ccaatgctgg	ccaagagcta	32340
caagcagga	caggtgctag	ggggatgtct	gaatgttcca	ccccaaagca	caggaataac	32400
ggaaatggag	actctaaagg	gcagaaagtg	agggtgtgca	gcaggggctg	cacaggacac	32460
atgcaaggcc	ctggctgcaa	taactgggtt	ggggagggcag	tcattgggcta	gccaggggca	32520
ccaggacagt	gtgccatcc	tgctccaaag	gcaggttcca	agccagattt	ctaggctcca	32580
gggggaggag	ggtccgggga	gaggggtcaa	gatttctccc	ctctgagtca	aggttggcct	32640
tcccatgtgc	cccaaatcag	gaggcacaga	aactgggatg	ttgtgggtct	acatccaagc	32700
tgagaagaca	agtgaggagc	agtacatgtg	tttcagatta	aaccagtcg	gagacaaaca	32760
tggtgtcct	cctcctccca	gagccaagct	gccttcaagc	cacatggcag	tgaatatgcg	32820
gacagtgcag	gggaggacac	ctctctctcc	actggctcaa	ggacagtttc	aaggggttca	32880
ggctggctgg	ctcatggcta	cgccgctcac	ccctgggaca	gtttgggggt	tttccctcct	32940
gaaatcttgg	aatctgaatc	agcctgagat	accccataat	tgtacctccc	aacaccccca	33000
gaaaggtcag	ccctgcagaa	cagaactctt	tggtcccccac	ccatccccct	cagccctgga	

ggctgaactg	atgggcagct	aaggtccaga	cagtggtgg	ctcttgaaa	gcctgtctct	33060
ttcctttgac	tcagaccact	ccctgccgtg	gcttacatca	ggaggtgcaa	gggctgcagg	33120
agggcagcca	gaccccaaa	accagctagg	ctaaatggtg	cttattgttc	gcaagaggcc	33180
atgacctcat	ttgtctccca	gctcttttgg	taagagagaa	tgagaggaag	ctggacagag	33240
aacctagcag	gcctcaggca	gccactgct	ccttgctgta	aggaaccag	caccgatgg	33300
tctgaaaagc	agcgatccga	atggagtcag	gctgagctgc	aggaagctca	ccttccttgc	33360
tcactgctgg	tggaaagcaac	ttcaggaaga	gcccagccta	tgggactata	gctcctccgg	33420
ggtactgctg	agtccagccc	cagagcttag	ctccctgctt	cccaccaccc	accaccacat	33480
cctttcccaa	caccattcaa	aacccagtc	cagcctctcc	tactggtcta	cagtgcagg	33540
ctaataagag	cctgggcctc	tgtccccca	attctctctc	ccctctcatc	tgttcacctt	33600
ggttcctaaa	ctgcaggggc	tactataacc	ctacctccac	ttccttgca	ccctcttttc	33660
tgctctctgg	ggtgccctg	ccactcccag	ccctctagc	caggagcct	cttccatata	33720
tgcttcccc	aggctagacc	aggcgtgcc	ttactgtgg	ttgaggcagc	ttctctcaca	33780
gcctgcactc	tgaggggctc	caggaagcag	tgaggggagt	agctgcctct	caaccagcgt	33840
ccagcaggct	tcagattaca	gctactcttt	tcttaaagt	acctgactcc	atttggaaat	33900
tgtgattgca	tcattgtctg	gtgttaactt	taaccactg	ctgcccttcc	gccatgtggc	33960
tccaagacca	cagttggcc	accctcctct	cccaccacat	ctcccttgga	tctttatctc	34020
tcttcattgg	gaccttcatt	gggacatgat	ggctaacttc	aggggcactt	gggccagcct	34080
ggggtaggtc	atgagtctga	acttgaacat	ctgaaaggat	tggctgagag	gcaggctgca	34140
tggagagact	gtgagccagc	cggataggag	atgctgggtt	cttccaggcg	cttggctctg	34200
gctcactgca	ggtgggagca	agggtattct	tctccctctc	tcacttgga	aatgaaggaa	34260
tgggactgta	cctgacagct	ctgaaggctc	caaaggacag	tggggtgggg	actagagagt	34320
tggccagctg	cttatgagca	ctggctgctc	ctgcagagga	cctgagtctc	gttcccagct	34380
cacatagcaa	ggactcgaaa	ctgcttgga	ctccagctcc	agagaatctg	acgctatctg	34440
ctnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34560
catgccttta	atcctagaac	tcaggaggca	aagcagatgg	acatctgaat	ttgaggcccg	34620
cattgtctac	atagcaagtt	ctaggctagc	caggtacatc	ataagaacct	ttctcaaaa	34680
ataaataggg	ccagttggca	aaatttagct	tgcctccta	acacaagaac	ccaaagtcaa	34740
ccccagcaac	catgtaaaaa	gaagcaaggt	gtggtggcac	ttgcttgta	tccagcattg	34800
tcaaggtgga	gacagacgga	tccatggggc	tcactggcca	gccagctagc	tggtctactt	34860
agtatgctcc	cagccagtga	gagactgaaa	aataaataaa	taaataaggg	gttaggaaga	34920
ggtaacatgg	tggctcagtg	agaaaagata	cttgtcatac	aagcctagca	accctcaatt	34980
ttcagtgggc	actaaagggtg	gaaggagaga	accaactcca	aagaattgtc	tcctgacagt	35040
tttatgctgt	ggtacacaca	cacgcacaca	cacacttggt	acatgcatat	gcatacaatt	35100
aataatttaa	aatgttatgt	gtatgggtgt	tttgcttaca	tgcatatctt	tctgtgcacc	35160
acatgtgtgc	aatgcctgtg	aaggctagaa	gaggacatca	gatccccctg	gagttacaca	35220
gggttgttag	ctaccatgtg	gattctggga	acaaaacctat	gggttttcca	aaagaggctc	35280
ttaatcactg	agccatctct	ccatcccctc	ataagtatat	atttctgggg	ctggagagat	35340
ggctcagagg	ttaagagtat	tgactgctct	tccagaggtt	ctgagttcaa	ttcccagcaa	35400
ccacatgggtg	gctcacaacc	atctgtaatg	gaatctgatg	ccctcttctg	ctgtgtctga	35460
agacagctgc	agtgtactca	tatacataaa	gtaaataaat	aaaccttttt	tttttttgtt	35520
ttgttttttt	tgtttttcga	gacagggaga	cagggtttct	ctgtatagcc	ctggctgtcc	35580
tgaactcgc	tctgtagacc	aggctggcct	tgaactcaga	aatccgcctg	cctctgcctc	35640
ccaagtgcctg	ggattcaggg	tttgccgac	cgccaccacc	tccaaggctg	ctgctgcggc	35700
caccaccacc	ccaccaccac	actacctgac	tatttaactt	ttaaaggcag	ccatctcatg	35760
gaaaatgaca	cctagcattg	tcctctggtc	cctacatgac	cccatgtgca	aacacatacc	35820
tgcataaaca	cacataaata	cataagtaaa	cttagtctgg	ttgttttgga	aatgtgctat	35880
ggttttgatt	gtgtccccc	aagggaaaaa	ttgggtccca	gagtagtact	attggaagag	35940
agtagaactg	ttagggttta	ggcctggtgg	gaagtggcca	tggctagaga	gacatgccaa	36000
ccaaggggaa	tctctggctt	catcttttcc	ctttgcttcc	aggtcctaag	acagtcacaa	36060
ggctgcttca	ccacatgccc	agaagcaagg	ggagcagtc	tggctgggac	cgctaatagca	36120
gctgttgatt	tccccagata	tttgtagtag	taagagacag	gtgaggaacc	ccacagcaag	36180
tgtagtaaat	tgtgtgtgga	ggtgccctcc	ggggacgggg	gccctcctgg	ggcaggacgt	36240
tcctcttctc	catccacctg	cactccgaga	acaggaaatg	gtgactttgg	cagagcttaa	36300
gcagagcccg	ttcatgttac	aagtatgtaa	attcataagg	accagtttct	ctccatatga	36360
aacagcttca	aacaggagaa	ggaagaagca	aacattaagg	aaaagctctt	ttattgcaga	36420
ggctacactg	aagctaccgg	ccgccttcct	ggaatgtata	atcagcttcc	ctctgggggt	36480
tctgtagagc	actgagacat	taagtactac	tggggtccag	gattctgcct	atgaagagga	36540
gggcccccg	gtccgtgtcc	ctcagaacaa	agaggaaagg	ttggttaagg	tgaagtcta	36600
gcggaagggt	gaggcggagc	ggctggaggc	ctgggctggg	gctgcttctc	gccccctctt	36660
cattccactc	gaaagcagcc	ctgtgttcca	cttggtgag	cttcacgggt	ttgccagtaa	36720
tcttgctgaa	gtcgggtgat	tcaaacaacg	actgtagctc	tgtggagatt	cagagattcc	36780
attaacacca	cacacacaca	cacacacaca	cacacacaca	cactccctgt	ttgtgtaggc	36840

tgattttcaa	gaaagcaagc	tagaagtgga	gtacctcaca	gtgacttgtg	agctatgagg	36900
cactctgtga	caggctcagt	gacctacctg	agaacttata	gccaaagatgg	ctgaagccag	36960
acctggcctg	agagaatgtt	ttgggctgtt	ataggacaca	tagagataca	cacacacaca	37020
acacacacac	acaccaagga	ctgagtcctaa	tgggaggtgg	ttcttcattc	ccctcccttg	37080
taatgggtgc	acatgttccc	tgagccaccc	tacaaagaaa	gccacaggac	tcagttctgt	37140
cagcaagggtg	gcaggctcca	agactcagcc	cogagcgcaa	agtggccttg	caaacatact	37200
catgtcctgc	agagacttgg	taagttcgcc	ttcgaagctc	agcttcagct	tggggacagt	37260
cagcacagct	tggatagtct	tcagttctcg	gtcgaagtca	tgaatgaact	cagaggtgag	37320
gctctcttct	atcatggtca	agttctgggt	cacggtcagg	ggcagggaaga	agatgatgct	37380
catacttctt	gtcaagggtca	gctgggcaat	ctaaccacac	agagatgcgc	acaggttagt	37440
tgtgagccag	aaaaacaaa	acaaacaaac	aaaaaacac	caacagctgc	cttccctct	37500
gctgtaacgg	ggccccagcc	ttgtgctccc	cagcctcagc	ctgggctgta	ggctactggg	37560
tactggcagt	ccctccatga	gtagggagtt	ttcttctcag	cctaaaaccc	acagaagttt	37620
aatgaacaca	cgtttgtttg	tggttccgct	acggtttcta	ttgtgataaa	acatgactga	37680
aagcaacttg	gagaggaaag	ggttttatttc	atctgacaat	tcgcagggtg	tcttctcatc	37740
actaagggga	ctcagggcag	gaactgaagc	ggaagccgtg	gaggaacgct	gctttctggc	37800
ttgtctcccg	tggcttctta	gcctgctttt	ttatgctatc	cagaaccact	tgcccaggag	37860
tgacactgcc	cattgtgggc	tgggcccccc	cacatcaatc	actaatctag	aaaatgacct	37920
acgggtttgc	ccagaggcca	gtctgggtggg	ggcattttat	caattgagtt	tcacccttcc	37980
aaatgactct	aacttgtgtc	aagttgacca	cacgaatcag	ggcctgggtc	ttaggagctg	38040
aagtggaaatg	ccccccagag	actgcctgcc	agcactgctg	accatttgct	ttgtatagag	38100
cattgaacca	gaaaatgaaca	ataaaatgga	tcctttgaac	agatgtgttg	atccttaggg	38160
ctgtggacac	agcgactggg	cttcccagag	ccccatgga	atcannnnnn	nnnnnnnnnn	38220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	38280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	38340
gagagagaga	gagagagaga	ggataaaaag	ttagcccagt	ggtgggtggc	cataccttta	38400
attccaacac	ttgagaggca	gaggcagggg	gagctctgtg	ggccagtttg	gtttacagag	38460
taagtttcag	aatagccagg	gctacacaga	gaaaccctgt	cttgaagaga	aacacacaca	38520
cacacacaca	cacacacaaa	taagatcttt	gggtgcgcag	aagaagaaaa	gaaaggatag	38580
tggggaaaca	tctgagcaga	ggaagaaatg	cacttgcaat	gacacccacc	ctcagaggag	38640
gccctcactg	gaggtgtctg	cacaggagaa	aggagcagtg	cagcttgccc	tagggcgctca	38700
gaactcagaa	ttcagtttca	aagcactgac	aaagccctgt	actggggacc	ccaggttgaa	38760
tccccctttt	atctaaaatg	agtaagaacc	aaaaaaacaa	aagtgtttgg	gatttggaat	38820
ctgggttatt	tgcatctaca	aaagaggtct	tggggaggaa	acccagttct	acccctggaa	38880
ttcatcagtt	tctataccg	ctgactacac	aggggctgaa	ggtaatctca	atgttttcat	38940
aactgggtgtg	gtgctacttg	ctcatgatcc	caacacttgg	aaggtaggtc	agaagttcaa	39000
gagcagtcct	gactactcag	tgaatttgag	gctagccctg	gctacatgaa	aactcataaa	39060
acaataaaaag	aaaagaaaag	gtggcgagtg	aggtggccca	tcaggtaaaa	gtgccaaacca	39120
cctcgctga	aagcctccac	acggaagggg	aaagccagct	tctacatgtg	gtcctctgcc	39180
ctccgcattg	accatggctc	gtgcaccccc	acacccacc	acccacccac	ccacatgaca	39240
taaataacttg	taatgattag	tttctgaaga	acaatatattt	cgttgatctt	gttttagggaa	39300
caaagttcgt	gcacattgac	ctgtcgaccg	tgtagtacgg	gatccgctcc	aggaagctaa	39360
agattttggg	atgcttcatg	ttgcagctac	tctgatgaac	agtgttgctt	tctgggccag	39420
gtaatgggtg	catatacctt	tgatcccagc	acttgggagg	cagaagcatg	tagatctctg	39480
tgagttccgag	atcagcctgg	tctacagagt	gagttccagg	atatccaagg	ctatacagaa	39540
aaacccctgt	ctctaaaaat	cactaattta	aaaaaaaaatt	cctttctaaa	cctatataaac	39600
aaatgttttg	taggctgcct	taacaaagcc	caatggccat	tcagagaagg	ctcaaaagag	39660
aaagtttagg	ggactctaca	agcatcctca	ggaaggccac	agaaagcaga	gcctggggcca	39720
gtgagacttt	gcagtgggca	aggttcagct	ctttatgtag	gaagaagaga	gtcaacagtc	39780
agagtcagc	tttccataaa	acctgtgcag	ggcctctagg	caaagccctg	tgttaggggc	39840
aaaggcattt	gcagtctaag	ccoggtgaca	tgagctcaat	ccttggaacc	caggtggaag	39900
gagtggtctg	actccacaaa	tttgtcctct	gatctataca	tgtatgcacg	tgcaacgaca	39960
ctcacataca	tgtccacatg	cacacatgca	tatacatgcg	catgcgtgca	cccacacaca	40020
gggtcaaaaag	cagcaagaga	tgccctgtga	aaaacgtctc	atcagctctc	ccatcatcca	40080
gtgccacact	ctgagcacag	gtgggtactga	tatcggttct	gattgatcga	tcagttgatt	40140
tgagaccccg	cctcactatg	tagcccaggc	tggcctggaa	ctcacagtga	tcctcttgct	40200
tctgcaagat	gagcccatca	tgcccacccat	gttattgaag	caataccatg	ctctataaag	40260
caaacctagg	caggcaggat	ggtggactcc	tgtaatctca	ggacttgaaa	agtagaaggg	40320
agatgaggag	ttcacatcaa	tctcccgtat	gcgttgagg	ctggagtggc	tgttcccttg	40380
gcgcttctgc	cagcacctga	ccaatgcaga	tcagatgct	cacagccaac	catcagatgc	40440
agctcgggag	cccagtgagg	gtgctggggg	gaggaactgga	ggagctgaga	cgggattgca	40500
agcccatag	aagaacaatg	tcagctggcc	aaaccaccca	gagctcccag	ggactagacc	40560
acgaaccgag	gactgcacat	gaagggatcc	atggctccag	atgcatatgc	agcagaggac	40620
agccttgtct	gacagcatgg	gaggggaggg	cattggctct	gtggaggttt	gatgccccag	40680

tggttgaggga	tgctggagcgg	gtggggcagg	agtgggtggg	taggtgggga	gcaccttcat	40740
agaggcaaaa	gggatggggg	agaaggcaga	tgggatgggg	gggttggtgga	ggggtaaaga	40800
agaaaaaaga	tgtctctgaa	agtaaaaagt	acttgtcact	aagcatgagg	atatgagtc	40860
accccccaagc	cccacagggg	ggaaggagag	aaatgagtc	cacaagttat	tttctgatct	40920
atacgtgcaa	tccatggcat	acgcagaaac	gcaaagacag	acaatgagtt	gggtgtgggtg	40980
gtgcacatgt	aattccatca	ttcaggagac	agaagcgaca	gagttgttgg	aaatctaagg	41040
ccaacctaaa	gacctacacc	caaagaagga	caaactataa	ggaaaaaggt	ggtcgaccaa	41100
tgtaacatta	aagttagaaa	tctctcttca	cactgtgtag	atactgtaca	aggaagagaa	41160
aaggcagcca	catcaaaa	gtgtaaatca	acgagaaaa	cagaaacaac	tcaagagaag	41220
gctgcagggg	cctgaattct	gttctcagaa	cctgcacaa	gccaaagaaa	tcaaaactgt	41280
ctgtaactcc	agctccctgg	gatccaacac	ccatttctgg	cctccatcag	catcactcac	41340
aggtgtgcac	acatacacat	caataaaaa	caaaaaccag	gatgaagggg	tagggaggtg	41400
catgtggatc	tgggaggagc	tgagaggcac	tgggtgaata	caataaaaa	tttgggtgat	41460
gggtgtgcac	gcctttaatc	ccagcacttg	ggaggcagag	gcaggcgaat	ttctgagttc	41520
aagaccagcc	tgggtctacag	agttagttcc	aggacagcca	ggtctacaca	gagaaacctt	41580
gtctcaaaaa	aacaaaacag	ccgggcggtg	gtggcacacg	cctttaatcc	cagcacttgg	41640
gaggcagagg	caggtggatt	tctgagttcg	agggcagcct	ggtctacaaa	gtgagttcca	41700
ggacagccag	ggctacacag	agaaacctg	tcttgaaata	aataagcatt	tgttgctgtt	41760
acagaaaact	ccagcccagt	ttccagcaca	cacagggtga	ctcacaacat	cataactcca	41820
cttccagggg	atccaatgcc	ttcttctgac	ctctgtgggc	accaggattg	catacagttc	41880
acagacatgc	acataggcaa	aacactcaca	aaataaaaata	aatctagcaa	aaaaaatctt	41940
aactaataat	ttaaagaaaa	aaataaggaa	gccgggggtg	gtgtcgcacg	cctttaatcc	42000
tagcacttgg	gagacagagg	caggcggatt	tctgagttcg	aggccagcct	ggtctacaaa	42060
agtgagttcc	aggacagcca	gggtacacac	gagaaacctt	gtcttgaaat	aaataaataa	42120
ataaaaaata	aggccaagta	attcttgtaa	gaatcccaag	gggacactaa	gtgtatataa	42180
aggcgttcca	tagggctagg	aatgaggctt	agcgagagca	acttcgctgg	tgtatgaaag	42240
tccctcagct	gcatgtggta	cctttaatct	aggtctctcc	gaagcagagg	cagaaggatt	42300
tctgtgagtt	caaggccagc	ctgggtgtaca	tagctagttc	caggacagaa	aggcgatata	42360
aatagaaa	tctaccta	agccngcca	aaanaagggg	agacctgaga	ccaagagat	42420
gactcagtg	ctaagagcat	tgactgctct	tccagaagtc	ntgagttcaa	ttcccagcta	42480
aaaatttatt	taaatgttta	ttacttgtat	tattattttaa	atttaaataa	ataagtaaat	42540
gggagcctag	gtttgagttc	ccaaatcacc	aagaaaaaat	gttatcattg	ctaataatca	42600
aattaagagc	ataagaactt	ctttttaaag	aattcttatt	tattttatgt	atgtaagaac	42660
actgtagctg	tcttcagaca	caccagaaga	gggcatttga	tcccattaca	gatggttgtg	42720
agccaccatg	tagttgctgg	gaattgacct	caggacctct	agaagagcag	tctgtgctct	42780
taagtactga	gccatctcta	cagctcttat	cagggtgata	aaatttaatc	tctgtggagcg	42840
ctgagaccaa	gaactaaagc	tgggagattg	aaaaatgcag	accaccaagg	ccctgctcat	42900
ttctccagtt	ctgatcagct	cccgtaccag	gggtctaacc	aggcctgtgt	ctgcttccct	42960
gagtagacca	gaggccccat	ctaaacagcc	tgctgtcagc	agctcctctc	tctaggtgga	43020
cagatgggga	tttcagacca	atgtcatttc	caggacatc	aacacagcag	ccaaatttat	43080
tggtgctgtg	gctgccacag	ttggtgtggc	aggatcaggg	gctggcattg	gcacagtgtc	43140
tgattattgg	ctatgccagg	aaccagtctc	tcaagcagca	gctcttctcc	tatgccatgc	43200
tgggggttgc	cctgtctgag	gccatgggac	tcttctgttt	gatggtcgcc	ttcctcatcc	43260
tcttcgccat	gtgaggctcc	ctggggctac	ccagccgtcc	ctgctgcctt	gactccatgc	43320
cagtcctggt	gctggagtct	actgagattt	caccattaaac	agcaacgttt	ctctaataa	43380
ctattaatta	attaattaat	cacgtgacaa	ccccagcgtc	catatgggtg	tggaaaatga	43440
ggaactctac	ccatcataca	tggcgactat	gaagaacaat	gtgacagaaa	atgctaacat	43500
catgtgtgac	cgcatgcac	agccctgact	gctaaaagtg	gacaagccc	aagcgaaagc	43560
ccaatgttct	acttctaaat	gcatgcacca	aacgcctccc	acaggaccag	aggtgcagct	43620
ctgatagggt	ccttgccctg	catgcatgaa	gctgtggaca	cgaggcatta	ttcgcaagaa	43680
cattctagct	gtctggagg	ccctcaatcc	actgtgttcg	cgctgttcca	gcaccagtgc	43740
ctcctggggc	tgcacctgaa	aaaggggact	gcttaagagg	gctcctacca	agcctactgc	43800
cacagatgca	tgatgggaaa	gccttctgga	agcaactggc	tgccaaaggc	tctggacaag	43860
agatcacccct	ctactggaaa	ggtgggttca	gtctaggttc	tgtgggattc	caggaaatta	43920
gacaacactg	gcagtccaac	agacagacga	tctaacttcc	caaggcacag	ctggtagaac	43980
ttgtctgcga	accagacaac	aaggtacgag	ctactccc	acaacataca	aaaaagcaga	44040
gagagagtca	gagacagaga	cacacagaga	gagacagaga	gagagtctaa	agagagtcag	44100
gggtctcagga	ctgagggtat	agtctactgt	agagcatttc	cccagcatat	acaagacctt	44160
ggattcaatc	tgaacacagg	aaaaaagggg	gggggggact	tcgattatct	caaattctcc	44220
tttttgtgac	acacccctaa	agtcaactgc	tacttccctc	accgcatatg	agtaaagagc	44280
tgtttgcgct	tatgtctaca	cagtcctggc	tcccacttcc	tcttcccttc	tgcttctgtg	44340
ctcatctcct	ctgaaaccac	tgagaaagtg	gacttgtgtt	gactgccaca	cggaaactct	44400
cctcagtagc	aggcagcaga	gcagagctct	gtcttctcgg	agcttcttct	ctcttgtcgc	44460
cattttctcc	caccttaag	tacctatct	tctctgtctc	tgcttgttga	tccttggaac	44520

cttttccttt	ctatgaacaa	aatatctcct	taaaggatct	cttctagttc	agggtccccc	44580
cgccccact	gtggagaaaa	cccagggcct	tgcacatgct	cagcaggagc	tccatccagt	44640
ctctagctcc	atgacttaaa	gcatctctgt	gctgtcaa	atacacttcc	agcccttacc	44700
aaaatattca	gtcaactcct	tgccattcaa	aatggatgac	ctcaaagcca	gagtcagcgg	44760
tgctatgact	cccagatcca	tccacttggg	agcccaggaa	tgaactcann	nnnnnnnnnn	44820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	44880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnng	ttactgggat	ttgaactcag	gacctttgga	44940
agagcagcca	gtgttcttaa	ctgctgagcc	atctctccag	ccaccaccac	caccaccacc	45000
accaccacca	ccaccaccac	caccccacca	ccaccccgcc	ccacctgctc	attcctgatt	45060
ggttgggttag	tttagtctgt	gagacaggag	ctgtcccttt	tctatagtgg	aaggtgaata	45120
agaaactcct	gaaagtgaag	gcctacaaaa	cagccacact	tatttggttg	aaaatactgt	45180
aaatgtgaca	tgtaaatata	tgctaaaata	attcggttaag	tcagtgaaca	accttaaaac	45240
acagtctgta	gcctgaatta	cagacacgac	acgagccatg	acagaggctg	aaataagacg	45300
cctttgcaag	gagaagggca	gaagcttcca	tcttgcttag	caaactcttg	ttccaagctt	45360
tatcagattt	tattgctttc	tctttctggt	tttctttatc	atatttggtt	atttggtggg	45420
ggaaagccta	atcttcatag	cccagtgtat	tgagcacttc	agcatatgtg	tgtgaacacc	45480
acagcacac	gtgtatgagc	accacagcat	aggtgtgtga	gtaccacagc	acatgtgtgt	45540
gagtacgata	gcacgtgtat	agagttcaga	ggagaactga	gagagtccgt	cttttcctcc	45600
tactgtgtag	gtctcagggg	tggaaacttg	gctcagcctt	ggtggcaagc	tcctttatcc	45660
acagagtcac	cctgccagcc	cagctttctc	ttttctcttc	tgttatgtct	atccactctg	45720
ttcaaggcta	actcactgac	tctgagttat	cagaactgct	tgtgagagca	ggagtaactt	45780
tggacatctg	tgtggtagg	aacaccatcc	ccactcggtc	tggatgacga	aggggaaaaa	45840
aagcatcacc	aaggagtctc	accacctcaa	ccagcaataa	tttacctcct	atacatggat	45900
aggtgggggtg	ggtgagcctt	gtgattttatc	gttaggatct	catgggagtg	attacagctg	45960
gtctactcca	tgaccaaaat	ggtgacgggtg	gctgacaaaa	aagaaacagc	tacacctggc	46020
tctagttttc	tttctttctt	ttttcttttt	cttttaccct	acggtactaa	ggattgaacc	46080
caggaatgca	agagctctgc	caagtgaact	acattcccag	atctgttttt	ccatttcttt	46140
ctttcctttt	agattttatt	togattttatt	tgtctatggt	tgtgtgtatg	tgtgtatgct	46200
tatgtgtacg	tatcagtatg	ccatgggtat	acagaaacct	gagaaggcca	gaagagagtg	46260
tctgggttac	aggagtttcg	agctgtcctg	tgggttctca	aatgcagcag	caccaccacc	46320
aatcaccctc	acccccaccc	agccttcgag	ttcaattctc	atcatcacia	aaacacacac	46380
acagacaagg	gcctgcaaga	tggctcagca	ggtaacgaag	ctcgtgtcat	aagcctgaga	46440
acctgagttc	actgtctgga	acccgcgtaa	agggggaagg	gaagaatcaa	ctctatgatg	46500
ttgtcctctg	ctctcccat	gtgtgccatg	gaatgaacag	ccctcccaaa	cacacatcaa	46560
gaataaataa	aactaaaatt	agcttagtaa	cttttatggt	gaaagtgggt	tttacatgcg	46620
tgggcaacaa	taacaccgag	agtagaaagg	caagcatgta	tgtcactgaa	cagcattgaa	46680
gaaaaaacaa	acacatttcc	tgtacatcgt	tctgggagtc	tgagttaggg	tttctatgct	46740
gggataaaaa	caccctgaca	aaaattaacc	tggggaggaa	gctgtttatt	tcagctttta	46800
tgtctacaac	atgacctgtc	acccagggga	gtcagggcag	aaattcaaac	aggtcagaag	46860
cctgtaggta	gaagctgata	cagaagccat	ggagctgctt	ctggcttgct	ctagcacaca	46920
ggagtactag	cctaggggct	gtactgcccc	cagtgggcta	ggctctccca	cagcaatcat	46980
aaattttagaa	aatgcactac	aggtttgac	acaggccaat	ctggtagggc	cattttctca	47040
attgaggttc	cttcttccaa	aaggacttta	gcttgcat	tgttgacata	aaaactagcc	47100
agcatattgg	gattatagat	attctcataa	aaaaaagaca	tttagattcc	cacataacac	47160
catattcaga	aattaactca	atgtgaacca	gaagctctga	aagtaagagt	taaaaactatg	47220
aaaaattctt	acaaccatcc	ataacaaaaa	tctgatgccc	tcttctggag	tgtctgaaga	47280
cagctacagt	gtacacacat	ataaataaat	aaataaatat	ttaaaaaaat	atatgaaaaa	47340
tcaggctggt	gagatggctc	agtgggtaag	agcaccgcag	tgctctttcg	aaagtccaga	47400
gttcaaatcc	cagcaaccac	atgggtggctc	agaaccatcc	gtaataagat	ctgactccct	47460
cttctggagt	gtctgaagac	agctacagt	tacttacata	taataaataa	ataaatctta	47520
aaaaaaaaaa	aaactatgaa	gaactatgaa	ctacaagaag	tcaggaatag	ggctgggggt	47580
gtaacccaac	agaaaaacac	ttgcctggcc	tgcgtttggt	ctctagcacc	accaacgtag	47640
aaagagaaca	gcagaggatg	agggcatcct	gacttgagtc	aagtgaacaag	tgataatcct	47700
cgagacacca	aatcacaa	gataaaagag	atcaacaagt	tgggctttat	ctgaataaag	47760
agctgtgtcg	ttaaatacca	cgcaggaagt	gaagaggagc	tgagtctggt	aacacaggcc	47820
tgaataccaa	gctactgggt	ggactgagg	aggacaacag	ctagctcaag	gccacctgg	47880
acgccagagt	taactcagag	agcagcttgg	gtagctttaa	tgagactctg	cccaggccag	47940
tgacacaggag	agatggctca	gtgggttaaga	gcattactcc	tcttgcaaa	gacctgagtt	48000
caattcccag	caccacagtg	ggcacttaca	atcatccata	actttagttt	caggggatcc	48060
aatgcccttt	tcacagtacc	agggcatgtac	acagtgcaat	tacatacata	catgcatgca	48120
tgcatacata	cacaggcaaa	acttacataa	aatacataagc	agataaatct	taaaaagaagc	48180
ggggcgtggg	ggcgcatgct	tttaatccca	gcacttggga	ggcagaggca	ggtgtatttc	48240
tgagttcgag	gtcagcatgg	tctacagagt	gagttccagg	acagccagga	ctacacagag	48300
aaaccctgtc	ttgaagaaaa	taaaaaaaa	aaagaaaaaa	atcttaaaag	aaaaggagag	48360

gactggagag	atggctccac	agttaagaac	acttgttctg	aggtctacag	agtgagttcc	48420
aggacagcca	ggactataca	gagaaaccct	gtttcgaaaa	acaaaaacca	aaacaacaac	48480
aacaacaaca	acaaaaccac	ttgttctttac	agaggacttt	ggtttgattc	tcagaatcca	48540
catgatgggt	cacaaccatc	agttgcaggg	atccaagggtc	ctgtcttctg	tgggcaccag	48600
gcataatatgt	ggtgtacata	catgtataca	ctcatataca	taaaataaaa	agtttttaaaa	48660
aggaggctgg	gtttgtagcg	cagaggtaga	ggtaaaaaga	ctctagcttg	tttaattgttg	48720
acatgaaaaa	aaaaagacat	ttagattcct	gcatacacacc	atatccaaaa	attaactcaa	48780
tgtgaatcat	aagctctgaa	agtaagaata	agcctagtat	gcactgtaag	gctctgggtt	48840
cactccccag	cactgcaaaa	gatcatgaaa	ccagaaatgc	agatcctctg	aaccacagca	48900
tgggaatgta	actcagccga	tgcagtgtct	acctgtcgta	tacagagcac	aggataaatt	48960
gatttgtgggtg	gtgcatacct	ataagctcac	tacgtggaaa	gtagaggcag	gacgaccaa	49020
ggttcagtgga	catccttggt	cacatagaga	atttgaggcc	agtctggtct	gctggtctat	49080
ttggaatgct	gtctcaataa	ataaaagaaa	gaaagaaaaa	gaaaagaaga	agtcctatga	49140
ttgtcttaac	ctctgacctc	tgtgttcata	aagctctcct	ctcaggaact	cactggtcat	49200
cttgtgaaaa	cctacccag	agtctctggt	cagaggaccc	aggctccagc	tgtggttacc	49260
acataggatt	tttatactag	aaaaataaaa	tgaataagta	tgtatttttt	aaaaagggtg	49320
agagctggat	atggtgggtg	ctagttatag	catccagaac	tgagacagga	tagccatgag	49380
gttgagaaca	gctagactat	acggtctcaa	caaacaaaag	taagggatct	gagtagatga	49440
ggttttaatt	tttttctttg	tgtttgttac	ctaactgtga	tgggtgtttt	gaatacatgc	49500
atgtctgtgt	atcacttgtg	tgcctgaaac	ccaagggaagc	cagaggaggg	catcgggtcc	49560
cccggaagta	ttattacaga	aggttgtgag	cagccatgtg	ggtgctggga	atcaaactctg	49620
aaagagccac	ctcgggctgg	agagatggct	cagtgggttaa	gagcactcaa	tggctgtctc	49680
tccagaggtt	tggagatcaa	atcccagcaa	ctacatgggtg	gctcacaacc	atatgtaatg	49740
ggatccgatg	ccctcttctg	gtgtgtctga	agacagctaa	agtgtactca	aataaataga	49800
tcaaaaaaga	aaaaagaaac	agccacctct	ccactctccc	tttttaaaat	cctcttgccct	49860
ctgtccctta	atgttaataa	cacagggtata	tgatactatg	ccttgtttat	gaatagaaaa	49920
tacacgtgct	aaagcaagtg	tgaaccttaa	atacattatg	ctgagtaaaa	ggagttagtt	49980
gcacacaaga	cttttctgct	caagagtatc	tgtatgaagt	attgaacatg	tgaactctga	50040
aatcgggagc	tgagggaagt	atggggagtg	ctaattggcta	caacatttct	ttttggaatg	50100
atgaggatgt	tctagaactc	aaaaatgggtg	ataactcagc	atataacta	aaactcattg	50160
aattgtacac	tttaaatgaa	tgcaataaaa	cttgtctcag	taatgtgggt	tagaagatgt	50220
acagacatgt	gtgtgtgtgt	gttaaaacat	ttcttggcat	ggcaataaaa	atacagtttt	50280
agccagggtg	ttgtggctca	aaaaataatg	ataataacaa	taataaaaaa	aatgaaaaaca	50340
gaggctggag	agatggctca	gcggttaaga	acactgactg	ctcttccaga	ggtcctgagt	50400
tcagttccca	gtaaccacat	ggtggttcac	agacatctgt	aatgggatct	gatgccctct	50460
tctgatgtgt	gtctggaac	agctacagtg	aaagtcaattg	caaggacttt	acaatagtga	50520
ccatgataac	attgaagcta	gacttgctac	tactgctgag	tgtgtctgct	ggctctttct	50580
aaaggagtaat	gttagctttt	tgtcttaaat	ttgtttcctt	cctttcctct	ctccctctgc	50640
tgttttttct	taccctctct	ttactttgct	ttccctctct	atctcctctc	ttaacagagt	50700
tgtcttatgc	agcccaaatg	ccatcttctc	gacctcagct	ccccagtgtt	gaaaaatact	50760
ctttccacag	gttatgttag	gagactggag	tctgtctcagt	cggggaggga	gcctgggtca	50820
agttctgagc	tcaattcctt	ttctttcttt	ctttctctct	ttctttcttt	ctttctctct	50880
ttctttcttt	ctttctttct	ttctttcttt	ctttctttct	ttctttcttt	taagacaggg	50940
ttcttttgta	taccctggct	gtcctggaac	tcactttgta	gctggcctgg	aattcagaaa	51000
tctgcctgcc	tctgcctccc	aagtgtctgg	attaaagggtg	tgaccacca	ctgcccagcc	51060
ctgggctcaa	ttcttaacat	tgtggagaga	aaagtattgt	agctgttctg	gccacctgga	51120
attactttgt	ttctgatctt	ttgtgtcagt	caaatccttc	tcatccatct	ttcctcgtca	51180
ggctataata	tagactctcc	ttgcaatact	tggaaatgct	ctacagtcat	ctacatcctc	51240
agtctgtctc	ctatattttt	tcctaagctt	ccttctaagg	tctttattgg	tttatgattt	51300
acacagaaca	tttttttttc	ttgtctatag	catgctgttag	agtgatcggt	gccagataga	51360
gaaagagaaa	atgagagaan	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	51420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	51480
cagctactga	ttctctctcc	tccctcctcc	ttctctctcc	ctccccagcc	tcatgctctg	51540
ctcatcttgg	acttctgcgc	atgtcctcag	cccagacctt	ctgctcttgc	ttctctctct	51600
cçcagcagcc	ccccagttct	cttctctgaa	cttctgaggt	actctccatc	acctctttg	51660
gctcctgctc	tgattgggtg	cacctgctgg	ataggcttgc	tcttgactcc	actgttcgtg	51720
tctcaattag	ggacctcac	cctctgatat	accacacatt	tccctagtgt	ctccacctcc	51780
cacccccacc	ctatacgcac	atacacactt	agctgcatca	ggatccctaca	ccagggactt	51840
cttacccttc	taatcctccc	caccggacac	tgcccaggga	cactggggct	ccagagggct	51900
attgccacac	ggacacacag	gagatctcat	caaggagatg	tgccctacccc	agagggtagc	51960
tctcaccatt	cacaagcaca	ccacttctgc	ctcagcttc	tactctctcg	caggaagtäg	52020
ccagcccggg	gccaagtatc	cccaactaca	tccccaaaat	tctcagacac	tgccagcctc	52080
cagctgtcag	cctggccccg	gctggcgggc	gcctgtctct	ggcatagcga	ctaggggtga	52140
attagaaacc	cgctagctcc	ctaattgcc	gttctgagct	gtccttggtta	ccggctgccc	52200

gaggcacaca	tagaggaaaa	ggctgagagc	tgagccaggc	tggcatggag	gtagccctag	52260
tagacctaga	gaggactggc	atgtggccag	ggaccaaacg	tggcacagag	agggctcagt	52320
gcaatctgcc	cogtgggtgc	ctcccagcca	catccatttg	cccagaactg	tgacgtcaaa	52380
ccagcccggc	ccattcattc	tttattcagg	tggcataaaa	atcactacaa	aaactttaca	52440
aaagagtctt	gggagctaaa	gggtcccttc	cttgccctcag	tccccaagat	tcctggcagg	52500
ggaggacaag	agagagaaga	aggaggaaga	ctcctggcag	tgttggcatc	tccaaatacc	52560
agaggggtga	cttgggtgac	aggacacagg	ttggggacct	gaatgtcttc	agcaagggac	52620
actctttag	ggtaggtcag	cctccaacca	tgaagtataa	caccaaggcc	agtctaagct	52680
tgggagacca	acacttgtct	ctccttttcc	cacccagggt	gtctggaaat	tgtctaaaga	52740
tggcctctcc	agcctctgct	tacaaatgtg	gagggaccct	aagttaggga	cttgcctaac	52800
ctacctctag	ccaaaaactgt	gtccacaagt	gccagcccac	aaaagatcac	cccctgagcc	52860
ccttgggaag	aaatgaagat	tcccatgcc	tgcccttcctc	caggccccac	cccacctgct	52920
gcaagagaac	agcttctaca	ctgggtgatg	tccttccggt	cccacccat	cccacaaagc	52980
tggttagaaa	gagtcacagg	agctgagagg	ctgatccagg	tggggactca	ggatgctgct	53040
gcccagggcc	cctcctcact	tgggggagct	gaactggggg	tagtcttctc	ccatgcgggg	53100
tgcaagtttc	aagtcaggac	caaaggtctt	gcctccatgg	aagtcagctt	tgtcattctg	53160
gcctatgagc	ctgttgtcag	gggaatctcg	ctgttctcgg	agctggggca	gcgcgtcggg	53220
gttaggttc	ctcacactgc	ccacaaagag	gggcacgcct	atgggtgtcct	ccatgatgaa	53280
gaagaggaag	ggtcggttca	cagtgaagga	ggagagggac	attcgattca	tggctacgct	53340
ggtagctgcg	gctgcctcca	caccagcctc	gctgagctcc	atggtagact	gatgttgac	53400
gctagacacc	accagattct	gctcagagat	cccacgaagg	tctgggccc	ggaacaattc	53460
ctgcaggcct	gcccagaaca	gcagatgact	ggtcagtgct	gcccacaggc	tatgtggatc	53520
tgtctagcat	cctggctaaa	gggaacactt	gaaccagcgg	gttgattgga	atctgttaga	53580
cctcagctca	gacaaactt	ctagaaacct	tttttttttt	tttttttttt	ttttaaatca	53640
ggatctgcgc	taggtacagg	acagaaagtc	tagaggagca	tatcaaagtc	tcccatccag	53700
gaagcagggc	cacctctggc	tcaggcacac	tggcagctcc	cgtactctgc	ccagaccacc	53760
taggggcacc	ctatccccaa	gctccttacc	cagttggctg	aggggtggca	ccaggtccag	53820
ctgctgttgc	agatggagtt	taggcagcca	caccttgggtg	ggcctctcct	gcagcgaggg	53880
atggttacaga	gtatcccagg	tcaggttggc	tagtacctcg	gacacgttcc	actcaaaata	53940
agtgggcatc	acgaccacaa	agctcatggt	gttcttaaa	gggaaatgag	ccacctacag	54000
ataagaaaag	gagagaacat	gaggaccaga	cagcacctgg	acctgtctgg	agctctgggc	54060
aaaattactt	ctgtactttt	gagacaagag	ccagaaattc	agggttagca	tgtcttctact	54120
taactgggtga	agtggaaata	taccacttac	ccctttgcaa	ggtgacatgg	gaccaaata	54180
gataatgctt	ttacacctct	ctgtgtgcac	acataagcat	atatgtttgt	atcgggtgta	54240
gtgtgtttgc	tcattgggtat	atggagtcag	aagtaggtaa	acatcagtcg	tcttctctaca	54300
ttgtctctcca	cttttttttt	tttttttttt	gtgttgccat	ctttttgttg	ttgttatttc	54360
aagacaggct	ttctctgtgt	agccctggct	gtcctggaac	tcactctgta	aatcaggctg	54420
gcctcgaact	tgcagagacc	cacctgcctc	tgccctcctga	gtgctgggat	ctaagatgtg	54480
tgtaactaca	catagctccc	tctttttttg	acacagggtc	tcattggatcc	caagctggct	54540
ttgaatggac	gttttggggc	tggagagatg	gctcagcggc	taagaacact	gactgtctct	54600
ccaaagggtcc	tgagttcaaa	tcccagcaac	cacatggtgg	ctcacaacca	tccgtaacaa	54660
gatctgactc	cctcttctgg	agtgtctgaa	gacagctaga	gtgtacttac	atgtaataaa	54720
taaattaatc	tttttttaaaa	agagaaagaa	atgatggcta	catacttctc	tctcgtctct	54780
ctgcccacag	tgtctgggatt	acagagctgt	acaacaagcc	caagtttgtt	gtgttttaga	54840
catgctaagt	tatcccaggc	tgtcctcaga	ctctctatgt	aattcagaac	gaccttgaa	54900
ttcttttaag	gtttattttt	atcttatgtg	tatgggtatt	ttgcctgagc	atttgtctgt	54960
gtaccgtgtc	cttgacgtac	cctcacagtc	cagaggaggg	caccatttcc	ccctgaactg	55020
gttgtgagct	gcatgggtggg	tgtctgggaat	caaaccctgg	tcctctgcaa	gagaagccag	55080
taagtactct	taactgctga	gccacttctc	caccttgagc	ttttcttctc	cctatctcga	55140
tctaaaagta	ctaggggatgg	cggatgtgcg	ttcatgtgcc	tggtttatgt	gttgctaagg	55200
gttgaacaaa	gggcttttgtg	catgccaggc	aagcactcaa	caactgagct	acacatcccg	55260
acagactttg	actcttctag	tagtagtgct	tccactacag	cctgagttct	ctatctgctg	55320
tcagcaagct	gtacaaacaa	gctatgggcc	ttcctgtcct	tgccctcctc	ttctctccgc	55380
aggtggggct	actggccttc	aaaatgaccc	atagaggagc	cacagcaaac	agtaggaagc	55440
ttgcccctcg	tcttttcccc	tctcccagag	agtcagctat	aattcgagtt	tttttttctc	55500
ctctctctct	ttaaacagga	tctggttatg	tggccctaac	tatcttcaac	ttcagttctc	55560
ctgcttcaac	cttctgagtg	ctgggattat	ggtgtaagcc	accacactca	gctcacacaa	55620
cctttttttt	tttttttttt	tttaagaat	ccatgcagtt	aggacagcat	ggaaatgacc	55680
aggctcaggg	ctccctgggt	accagcataa	tgcttgccag	cgggtcctct	gccagtgagg	55740
ggatggaaa	atggagccag	aggatctttc	ctctctgaac	ctcaatgtcc	cacagtgaga	55800
cactcatgtc	cactgggaga	tactgtagta	ttcaaggaag	aagcaacagg	aaggtgagag	55860
ctaagtggag	ctgagcaggc	tcgtatcctc	tcaccacggg	ctacagagaa	gtctgtgctg	55920
cccctccaca	tggctcctcc	ctgcagaact	ggcaatgctg	ggccccgctt	gccagtcac	55980
actaaccaac	agaatggatg	agcatgtgtg	gtgccacaca	cctgggaccc	cagcactcag	56040

acagctgggg	cagaaggggtc	atgagtccaa	agcgaacttg	tgtaacattg	tcagaccctc	56100
gaacaaacaa	aactagcccg	tcctgtttatc	tcagccacag	atgatggggc	caaggatcag	56160
tactctagcc	aaggagtcac	ggttaggcta	gaagcaaggg	aagccttagc	tgagacagct	56220
tgggcagggag	cttcatccaa	tcagaatggt	cagagcaata	agctttgaaa	cccgaacttc	56280
atctatgaag	cactgtgtgg	gaactcctct	cttcctttac	gagcagggcc	ctggctcctc	56340
tggggtccgc	taaaacccca	gcacagagaa	cagttacctg	gcacgtgaca	aaaactcaat	56400
atatTTTTctt	tgaggagatg	aacctcaaag	aagctgtgtc	ctggatagac	acagcataat	56460
aaacccttca	ggagctacct	acccagggac	cagactttac	ctcccagtac	caggcctcgt	56520
ttgccagcca	aaggcaaaagt	ccagactgac	ctgtatctca	ggttgctcca	gcaggaacca	56580
tcgaagagga	tatgacaccg	cgtgcatcat	gtccaccgac	actgtgaacc	gctcatccag	56640
gtggaagaaa	tctttctggg	tgaggctcgg	gtcaaaacttg	gtcctccaga	aacctgcagc	56700
caggcagagg	gcaggagcca	tgtaacataa	aatcagcctn	ctgcctgtct	tgccatagaac	56760
ctatnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	56820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	56880
ggaaaaagga	atcttaaatt	agaagatgcc	ttgataagat	tgccatgtag	gtatgtctca	56940
ctaattgattg	atgtggaaag	tcacgagggga	tggtgtcacc	ctgggcagat	ggcctggggg	57000
atataaaaaac	acaggctgaa	caaaccacaa	agcagtagtc	ctcaatggct	tctgctttag	57060
tttctgtctc	aggttcctac	cttgacttcc	ctcagtgaa	gcatgtcaca	tgagagtgtg	57120
aagaggaaat	aaaccctttc	ctccccacat	agtttttggg	tatgatgtta	tatgtcaaca	57180
acagaaacta	taactaatat	agttggtttt	ctttttttgt	ttgttttgtt	ttgttttgag	57240
acagggtttc	tctgtatggc	cctggctgtc	ctggaactca	ctttgttagac	caggctggcc	57300
tcgaactcag	aaatccacct	gcctctgcct	ctgcctccca	agtgggtggga	ttaaaggcat	57360
gcgccacctg	tgccctggctg	gttttctttt	ttttttaata	catttataat	gcattttaga	57420
tttaaaaaaa	aaaatggcca	tgccatataa	tataaaaaaga	agtgtttaca	aatcaccatg	57480
tgcccttgcc	ataaattatg	taaaaatttc	catatggaca	tcagtctcaa	gcttacaatc	57540
tcagcactca	tgagcctgag	gcagaggcag	gaggatgggt	agctcaaggc	cagcttagtc	57600
tacataacaa	gatcctgtcc	aaataataac	aacagtaata	atttcataca	tagaactaga	57660
agggggccact	gcaaagacag	tatgacaaaa	ccactggccc	tgccctaattg	tattttaaat	57720
aactgtcctc	ctctctgtaa	ttttcagttt	ctaattttta	cataactacc	atgtattctt	57780
tttgtaattt	taattagttt	tttaataata	gaaacaagct	aagtgtctaag	aatattttca	57840
tatgaacatt	ttcaaggcac	ttgatacata	cctcagattt	gccctccagg	tgagcagtac	57900
caattacgtg	ccaccagcaa	tgtttagcttc	cttttttccc	taccatctga	ttctgtttca	57960
gtctattcgt	agttctgate	ttgttatatc	cctttttatt	gtttccctgg	gttccaacac	58020
ctcccagttg	agtgtttcta	ttgaatttca	tttagcagctg	tttcattaat	ggcacagaag	58080
aaggattaca	gtgttaacta	ggatagactt	tgacaaagaa	ctatgagaac	atatcttatt	58140
atctttgcat	aaattctttt	taatcaaagt	tctcaaaaag	cctctctctg	ttcccatctc	58200
agggagtagg	tctggccact	gatgagtgtc	caggccacag	tacagggtgtg	cgtgggttctg	58260
tcctctgtgg	aagggcacat	ctgtgttgta	acaggattcc	tgtcttaaca	agccttgctc	58320
aggtctctaag	tggtcctgag	ctagctaact	gcccttggtg	ttcccttgat	taccagataa	58380
ctattcactc	ttctcatttt	gcagagcact	taccaggtag	ctatgtcctg	gaagtacgaa	58440
tgagtccttc	tattgttttt	cttttactta	aatcccatct	gaaatgcgcc	agggacactt	58500
caatccaagg	tacacttttg	ctaaagaatc	actcattttt	atatgcaaaa	tgtcacctat	58560
taactgcagc	tgatatggta	catacatatt	ctctcttcct	attatccact	aatagggtgac	58620
taatgcgaaa	tattgagtaa	tttttaaaaa	tcaatactca	attttttaga	aataattaga	58680
gagacattca	actctgacac	cagcacctca	ctcagttcct	gagccttcct	ctgcccaggg	58740
agaatctata	aataactcac	gaagctgaca	ttactcactg	tggtgcagtc	atttttttct	58800
gagaaaattt	tagcaactgt	tctaataagag	cctgccagtt	atcagtagtt	gagaatgcaa	58860
gtcaactttt	aattatgcag	acgctgatta	ttcagacgac	aaattgtttg	tgccctgcacg	58920
gtcctctcct	gctgcctacc	tttaaccggt	ctcagtgctc	attagcacat	gttccagaag	58980
gtaggctttg	gaggggcgga	caggcactca	aaccagctaa	gcacttagag	aagctctgat	59040
gaaagatgtt	aatgcagttt	gtagaattat	tgactaaaaat	tgagtcattt	ggattccctg	59100
tgaattgtat	ttacatgccc	tgtccctgtc	ccccatagca	acagataata	ggattgtctg	59160
cagagagaca	acatagttct	tatattttaa	tttttctttt	gtcgaacatt	ttcacatgat	59220
ggttcgtggt	gtttcctttg	ttcattacat	ttgtatccag	actagttact	tctgataagc	59280
ggtttagttag	gattcctggc	acgcggacag	tgacaccaca	gttgtctgat	cgtttccac	59340
ttttttacaa	aaccgttttg	cttttaagagt	cagtgttttg	cacatttcac	ccagattatt	59400
ggaaatatta	tttccctcct	gcttaaacgg	aagctgtgat	cataatttaa	gcctttctag	59460
gtagccgatc	ttacatgtat	catacctatt	tctggcatat	gtttgtctat	tacaaagacc	59520
tcgtaggtat	gcagtttagaa	gcctctagtt	aaatgaaatg	ttgcgtgtgt	gatgaacctg	59580
gagtggggat	ggccttttgt	gtgccccaa	gctgttgtgt	ttcacacagt	tgttttctgc	59640
ctcctctggt	ctatcactat	cctgccactg	ccagaaaaac	ctgctgtgtg	ttcccgcgt	59700
ggaggatctc	tgcttctgaa	cttctttggc	ctgagaaact	ccataaccac	atcagtttagc	59760
attttgttta	aagagcaggt	aggctgttag	agcttgggtc	ttacatgtct	cccaggtcca	59820
cttgccagcg	ccttgaccac	tgtaaacctt	tgtaaaccaa	ctcatctttt	gctgcctggt	59880

ttttgggggg	tttttttggg	tttgtttaag	ccaagatcag	ttatatggcc	caggctgagc	59940
ctctcttccc	agcctctcaa	atgtttagaat	tacaagcatg	catccctcag	catacctttc	60000
ctttgctttt	tttaaaatag	agttttgcca	tagcaacaga	aatctaacct	aactaagcat	60060
agccgtgcac	atgggtatgag	gaactcacat	atgtgtgaat	ggaagtccat	agagaccggc	60120
atcactgcct	agaggccctt	ttcttccttc	cttgcagttg	tcgtgctagc	tgactgtact	60180
acaaaagagg	ttgtctgagg	cataagacta	ccttcaataa	aacatgcaca	gacagtttgc	60240
ttctctgaga	tttcagagca	gtgactacct	tcaataaaac	atggacagac	ggtttgctta	60300
cctgagactg	cagagcagtt	tccaaaaatt	ttagacaaag	ggtaggatga	agaaggctgc	60360
ggggttttgc	acacacttaa	gggtgcgtaag	taaataaaact	gagctacact	gacaggatgc	60420
tcgttctagt	agccaaccaa	agagcagttg	aaccaaagca	cctagacttc	aaacatcgtg	60480
gggagataat	cttaggagtg	ctatgcttct	gcgtccctaca	agtattatga	aactgtctag	60540
aaagcaccct	actggtaatc	cctttttgat	tatttttttt	ataaattcta	gtcttggggg	60600
tttgagtggc	acacagacat	aatgggttagg	cttcgggtgtg	tgctcattca	ctttgcttcc	60660
tggggaccag	agttttcgat	gagtcagtgt	ccatctgatt	tctgtcggat	ccggctgcag	60720
agccatgact	cagatgggct	tcaggcccag	ctgctcagtt	catcttctgg	ggaatagatg	60780
acaaggacgg	gacaaatgtc	ctgacgcaca	ttctcttctg	ttcttgcaact	tccagggtct	60840
aacgagagca	tcattaccac	cagcaggcag	atacgccttg	ccaaggcat	cttccctgtt	60900
gtcagcctcc	tgaaccactc	ctgcaggccc	aacaccagtg	tgtccttcac	tggcactgtc	60960
gccaccgtcc	gggcagcaca	gaggatcgca	aaaggacagg	agattctgca	ctgctatggt	61020
gagccagcct	ttctttccac	taccctgctg	tgccctcacac	ctcacatgaa	aaggataagg	61080
ggacaggaat	cagcagatat	gggcccagtg	cctctactca	tcctctgagt	ctttctgga	61140
aaggggcaatg	catccttggg	ccaataaaaa	aggtcttctg	gctgtaataa	aaaagcccgt	61200
tgagggcagt	gagccatatc	cctccatgcc	ttgtagacag	cctatcctga	aaatgagcga	61260
ggagcacttt	cttggcttct	ttcttctctg	cccagcagct	tggaacgta	tccactttca	61320
cccgtgtttt	gttgtttttt	ctgagatgat	agggcagagt	acccaacctc	atataggcta	61380
ggctagtgtc	tatcactgag	ccaggacccc	aaccagcac	caccatgcca	gtcacgtgat	61440
gactaggcca	gcccctcggg	agagtaggca	ttgactctct	tggtgtgact	aggaactgtg	61500
ggtaatctct	ctccagggcc	tcacgagagc	cggatgggcg	ttgctgagag	gcagcagagg	61560
ctgagtctct	agtacttctt	tgactgccgc	tgtggggcct	gtcacgctga	gacactgaga	61620
gcagctgcag	ctcccagatg	ggaagccttc	tgttgtaaga	cttgacagag	gctcatgcag	61680
gtaaatctct	gctgttccca	ggggcagggc	tccagctaaa	ggttgtcagt	cgccaggaga	61740
accattctctg	cttcccttct	tgtaactcct	ccctacatgt	cgcccgttcc	tgcaaaaaac	61800
acaggttgta	tttccctaata	ttttccctat	aagtgcacaca	aaatcttaaa	ttacacaaaag	61860
ggaccacaaa	aaaaaaaaaa	aaaaaagccc	tagaaattta	cttgctcaaa	taagtcatca	61920
aaagttgtgc	atcaggccta	gcacttgggt	actggttaacc	ctagcactca	ggaggctgag	61980
gaagaaggat	ctcaagtcgg	aggccagctc	caagtgcac	cccattctaag	agatcaccat	62040
tccaaggagc	tatttcagag	atggtttaat	ctggggaccc	agattgtgga	ttttctgtct	62100
gttcaattcc	atctctctgt	gctggcctca	tcagacacac	tctgtagtaa	ctgtgggaaa	62160
atccgacca	catagttttc	cctcagcctt	tgaccagag	ggaagagcca	cagtggagag	62220
catgagagca	gacccttggg	tgctactgcc	aggtaatggt	gtagacactg	gagtcttcaa	62280
cattcatgcc	ccaatgcaaa	atggtctcca	caccagagca	tggcattctc	attagaaata	62340
agtaaatgga	attggctgtg	ttgaaaattg	taaagccaag	ggtcaagaat	gaagccttcc	62400
ccagcatgtt	ttgttttgtt	ttgtgtttta	ggcagcgtct	ctctgtgtag	ccttgggtcc	62460
tgccctctgc	tacctctccc	aggtgtgcca	ccatgctggg	cctaagcgcc	ctgtgcatta	62520
gtgctccctc	gatcctgctc	actcttgaga	cagcttctct	tctactctgt	atccccagat	62580
aacctagagt	tcacttcaga	gcccaggctg	gcctcaaact	tgagatcctc	gtgtcccagc	62640
ttctcaaatg	cagtgatatt	tacaggccta	cacctggctt	tccctgatag	attcctagta	62700
agatgattat	cctttgagcc	atatctctct	tctgcttctt	cctctcttcc	tgagggttg	62760
atctagaatt	tattctaaag	ctgactggcc	tcagaattgc	catccttctg	cctttagnnn	62820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63000
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63060
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63120
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63180
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63720

cctttaaccc	cagcactcag	gaggcagagg	caggcagatt	tctgaattcg	aaggcagcct	63780
ggtctacaga	gtgagttcca	ggacagccag	ggctatacag	agaaaccctg	tctcaaaaaa	63840
aaaaaaaaaa	aaggacttta	aattgggctg	gagagatgga	ttaaaagcat	tggctgtctct	63900
tcccagaggt	cctgggttca	attcccagca	ctcaaattggt	ggctcacaaac	tgtctatatc	63960
aacgcaatct	aacaccctct	tcaggcatgc	aggttacatg	tagacaaaac	atccatatgc	64020
ataaaataca	taagtaaata	agtcttttaa	gtgtatactag	aagctgggtg	gtggtgcatg	64080
cctttaatcc	cagcacttgg	gaggcagagg	cagggtggatc	tctgagttgg	aggccagcct	64140
ggtctgagta	aatagagcct	tgtacttcta	cttatcacta	cagttacatt	ttataacttt	64200
gggcccctagt	gcttccattt	tccactgttt	gcttaaccac	tggggcctga	agctttttgtg	64260
ctgacacttt	tgttcgctaa	tcacacaggca	accaatgggtc	tctacactcc	atcacatca	64320
acacaaacaa	aacaaaacac	aacactacgg	atccctggcat	ggtggaacat	ctttagcccc	64380
agtacgtggg	cttgagttca	aggccggcct	ggtctacata	gcaagttcta	ggatagtagg	64440
gatagtcttt	aaaacaaaac	actattttat	ttatgaacaa	aacatgtaaa	gaaagaaaaa	64500
aaactgcaaa	tttatctatg	aatgaagtct	aagtaatact	tcaatattgg	aaatagcttt	64560
ctaaaatatt	tttattttaa	gaaaactcag	caaattattc	aaacaacctt	ataaacgttc	64620
gttataaaag	taaagaatta	tttgcaattg	ccttaagggt	ccaagggtggc	agcctcttaa	64680
aattcagaac	aatccaagct	tcacattcca	gttcaacatt	tctacagccc	taacgtattc	64740
aaatacctcc	attctgacaa	ctgtttcccc	tcttcttttc	ttctaagctg	cttagatgtc	64800
tgtcccaggc	ttttcatgat	tttagtcatt	cacacaacta	gcaaacatta	tctagggact	64860
aaaacttgcc	agatactggg	atatcaccct	aaagggggac	tgaagtagtc	tgcaaggctac	64920
agtctctaca	atctcctgaa	tgaatacaaa	agtagctaata	atttaccaaa	taaacatgta	64980
cacctgtgat	gattgctagc	tgtactagca	gaagctaaac	actaaatcta	gaaactcagt	65040
cctccaacta	tccccttgct	cggcttcagc	ctcattttta	caaacaaggg	aaagagtttg	65100
gaatgttgcc	caaagccata	cataagtga	caaaaaggag	ttggagtctc	caaatgcatg	65160
gatttgggct	agttactttg	ccaaccaact	cagtaacaac	tgagctgaac	aggaacactg	65220
tggtagcaaa	agaaaactgga	actatcaatg	gcctctagag	caaaaatata	tttaaaaaga	65280
aaaaaacaaa	caaggcctgg	caaggagact	gtgagaagag	tgtgctgact	gaaattgact	65340
agttcagcca	acaaaagact	attccagggc	tggtgagatg	gctcagtggtg	taagagcacc	65400
cgactgctct	tccgaaggtc	aggagttcaa	atcccagcaa	ccacatgggtg	gctcacacc	65460
atccgtaaca	agatctgact	ccctcttctg	gagtgtatga	agacagctac	agtgacttta	65520
catataatca	ataaataaat	ctttaaaaaa	aaaaaagact	attccagtggtg	ggatggaaaa	65580
gttaagtgtg	gagttaaaaat	atacttcaac	tggtgatgga	ctaggtgtcc	agagtccggc	65640
aaaaggatgc	tctgtggtag	aggtgcctgc	tgtgtaagcc	cagctacctg	agctcaatcc	65700
acagaatcca	cagcggagtg	ggaagagaaa	caacgtccca	gagttgtcct	ctggcatccg	65760
acgcacattc	gccatcccca	agatgtcata	catatgtgta	catactacac	actggcgcac	65820
gcgcacacac	actctttttt	aaaattcaga	cttagaggga	cataaaggat	ttgctctgat	65880
atatgttcaa	ttgaaaatga	ctttgaagat	agagggcaga	tcgaaaggaag	ctcagcagga	65940
aagaattaat	aacatgcagg	tgaagggtca	taaactagtc	tgcaaggggc	cttggctcga	66000
caaaaaaat	tatggggttt	gccggtaaaa	taaggaaaaa	gttgtcaaca	tgaaacacag	66060
aacactagca	agagaggagt	gttagcagaa	agaagccaac	aagctcaaac	aattaggtcg	66120
gctgaaaaat	tttaaaatgt	cttctgattt	ggctactggg	aagccactgg	tgacttcggt	66180
cagcgttttc	tctctcgtga	ccagagagat	gtctagtagc	aataatgagt	taggaggtatg	66240
taaaagaagt	aaaacagccg	aaaacaagtc	caaaaagttt	ggggtgatgg	agaaaaggag	66300
gaaacagagg	ccgccgaaga	tagacagcgg	catgtttatt	tgtcttgttt	tcttagatgt	66360
aaacaaacta	aaaaaactcg	tgagttcttc	tgcagtagcc	gggttgccctc	cagcatcctc	66420
tgatggtctt	agagaccccc	ggatgctccc	ccgcggccgt	ataatttcct	ccctgacgct	66480
ctcccgatcg	acagcggtct	cctccccggg	tcctctttgc	accgctccaa	ggccgcgctg	66540
ctagggccat	cgagccccgt	cagggtcgtc	tccttaacctc	gatggccccc	tcgctcaggt	66600
gtcccaccat	ggctgcaccg	ctaaactccc	cgctcgcgct	cttgacaccg	ctgagcttct	66660
ctgcgggggt	cccgcgggct	gctcaacgat	tggctagagc	aactgtgcgt	gccgatccgc	66720
ccccagcgtg	agcgcgggtg	gaggggcggg	cctagacgcc	gatagccacc	gcattggcta	66780
ccgcgcggca	ggcagagcac	gtgactcttc	cgaggccggg	ttcgaggcct	agtggcggga	66840
tggcgggacg	tgagggcggg	gcgctgggtc	gcagtgcgcc	tgtgtcagcg	cggtgtact	66900
gagttgttcc	cccgccagct	gtcggaaact	tgcccgccca	gtcctttggc	ggacagacag	66960
aatggcaacc	cagggaacag	tcggagctct	cccctggtaa	ctgctgctaa	atatagtcaa	67020
agcagtgacc	tgggtacttc	ttcacgcagt	gcgtgcccgg	cgccgggtgcc	aggccagag	67080
cttggcactg	tgggataaac	aaggtaaatc	agactcagtc	tccgcccctct	tgagttccac	67140
ctgagagttg	tggccgcaag	gaaccagacc	tcaaggatgg	tagacgcgat	atgggccaca	67200
catgtggagc	tccagagtggt	gggtcaaaaa	tcaatcaggc	tttcgagagg	cgatgcgggt	67260
tgaactgagt	taaagtgtgt	gtagaaattt	gtcaggtgga	ttccagtggg	gatagtgatg	67320
ttcctaaaag	cccaaatggc	ctatgggaga	gtattgggag	gcctggcggtg	ctggctggct	67380
ctgatctggt	tgtaatccca	gcctttggga	tgtagaagca	gcaaaagtcc	aaggtcaccc	67440
ttaacaccgt	tgagttcgag	gtcaacctga	actaaatgag	accctgaaaa	atcaaaattt	67500
gggacccagg	cgtggtggca	ttcagaggtaa	aagcaggcag	atctctgagt	tcgaagccag	67560

ccaggctaac	ataagatccg	gtctcaaaaa	aaaaaagtaa	taaaaataaa	aaggaggaga	67620
ggctatatga	actgaaagaa	agacctggag	atcaaaacag	aaaactgagc	cgtctaagaa	67680
atgaaaatat	ttaacttcat	agttgctgga	gtaagaagtc	tggaaaactt	tgggcaacta	67740
aggtaaacag	gtctagaaag	actggaatag	tagccatcta	ctggatattt	gatctctgtt	67800
tgtacaacca	caacctacta	tagttttctc	aacaggtcca	aagaatatgt	ctgggtgaat	67860
tggtagccaca	ccacagatta	actctccttc	agcatatcaa	cagctataga	aaaccccgaga	67920
agaaatgatt	ttggttgctg	gtcacttggt	aggatgaaat	ctcgattttc	tagaactatg	67980
cattaataga	aagctgaatc	ttcatgttct	gactttacag	agctgcggca	gcatggatct	68040
accggtggat	gaatggaaat	cctacctact	taagaagtgg	gcttcaactc	cgaagtctgt	68100
gcaggacaca	atttctacag	cagagacttt	gagcgacatc	ttccttccct	cttcttccct	68160
tcttcagtaa	gtgaatggaa	acttcaggga	aattttggtc	tggaaaatgt	tctgccttgt	68220
catttggtct	gaatatctct	tttttatagg	agagagtagc	tttatattct	ttatagtatg	68280
gggcatthag	cagttactgt	tggttttcac	gtttctccct	agtctgtgat	tactagaatg	68340
ggtaggcact	aactgctttc	ctcttttggt	atgtgttata	cttaagggaat	gtagtatctt	68400
gctgtcgtcc	cagtgtctgc	actcatagga	tctggtgcag	gttgtgtagc	tgcccctaga	68460
agctcaatca	gtcctaattg	ggagaaagaa	ccctggcact	tggtagttg	agaccanaaa	68520
cttctcaagt	gtcnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	68580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnntttagtt	68640
tccaagtgcc	actttactgc	aatgtgtcac	cacatccaga	gttctgtgtt	tgtttatttg	68700
tctgtttttg	agacagggtt	tctctgtgta	gccctggctg	tcctggaaact	cactctgtag	68760
accaggctac	cctcaaattc	actgagatct	gcctgcctct	gcctccagag	tgtctgtgta	68820
cactaccacc	acccatccag	ctttctatat	tggttttcct	atggcgtttt	aaaatagcca	68880
ttatacgtgt	gtttatcatc	taaagtcctg	gtcccaaaag	gagatgagag	gggctgctaa	68940
ggtagaaaag	attacaaacg	cttatcaatt	ctgttcaaaa	attaaacctc	agagtgggat	69000
tagctgcttc	tttcattaga	attgctatca	gaattcactc	aggccttggt	tgctgtgtgt	69060
attgaagaag	tctcttcctc	atcagggtcag	tgactcctta	gctcaagtac	atgcaatatg	69120
cagtattgat	aactgttctc	gcttaggaat	aaaaatagaa	ctgacttcca	cagggaatg	69180
atgtgctgag	ctgtagcaac	gaatcttgca	caaaactctg	cagcagggac	cagctagtct	69240
ctcgctgca	ggaccttcag	caacaggctc	gcatggccca	gaagcttctc	cgaaccggta	69300
aaccagggtga	gattggctcc	ctcgccctag	gcctcagccc	ttcccttggt	tattttggta	69360
tcaccttgcc	ttactgagca	gtcctcaata	aatgactgag	gacttgaatt	taattatccc	69420
agcaccagcc	acaagatggc	tatgtaggcc	agttagacca	gactgtgacc	agctgttact	69480
ctgggtgccct	tgaagtcctt	cctgatgggt	taagctgtgt	ctgctgcgcc	agatagttct	69540
agcagctcga	gcaccagaaa	ggctgtctga	cttctatggg	ctttgtgtgg	ctccaggagt	69600
ccaatgccat	catctgattc	ccagcttaag	gacctaagct	ccgagaaggt	tgtctgtccc	69660
tcagcagcag	cagcaagtcc	tgagtgtctg	ctgggctcgt	ggtgtgactc	aggagttagag	69720
ctcggtagct	agcctgagct	gagagctgag	agaaagaaa	gactcctctc	ttttcagaaa	69780
gggatttgca	gaactcgatg	ttagaccctg	acatggtagg	aatctgtttt	gactattcta	69840
gcctagattc	tgaagttgac	ctttagccta	gagtcagaaa	aactaatgat	tacaggagga	69900
atgtgtagtt	tgttggttaa	tgttggttgg	aaaatggatg	ttagaagccc	agggtaaatg	69960
tgaggaagcc	tcatctaaca	cctcttttac	tgaagagaaa	aacataagca	accaacagct	70020
tccctggaat	gcccggtgtg	tgactccgtg	agataaagag	gcattttcac	tttgacctaa	70080
ccgatagaga	ccttgcaacg	tggctctctg	tgtccaggac	tagatctgta	tctgttgtga	70140
ggcatttttc	ctttgaaatcc	atagagcaag	ccattcagca	gttgttgccg	tgccgggagg	70200
ctgtctgagc	cttcttgtca	gcagagcaca	ccttcatggg	ggaaattgaa	gatggcctgg	70260
cccaggccca	tgctacctta	ggtatgtctc	cttaggtata	gccggagtct	tccttccctg	70320
ccgtgtgttc	agtgcggccc	ttgccttgct	tgtttggttc	tctcttgcca	tctgaattga	70380
cgctcttctc	cctcccatcc	tgcatctcct	gccccagag	ccttaggcta	atggtgtttc	70440
ttttccggaa	tgagacattt	ctcttctcac	agggaactgg	ctaaagtctg	ctgcccatgt	70500
acagaagagt	ctccagggtg	ttgaaactcg	ccatgggcca	tccagtgttg	aaattggcca	70560
tgagctcttc	aaactggccc	aagtccctatt	caatgggtag	gcctttcttt	ttcctagtgt	70620
ttggccaggg	cacacagtgc	tctgtgtttt	cctaggtgct	tctgtgtatg	gctttttgct	70680
acagtgtctt	aaagcatgtt	gaaactcttt	tatttccctc	ttaggacaga	tatttgccct	70740
ctgcttcaact	gatagacttt	aagctttgaa	ttccttccct	aggatgtgga	gaaagccatt	70800
aggtctgcat	ggagcttccc	agggaggatt	tggaggcagc	ctcaccgcgc	tctagcattc	70860
ctgtctgctt	aatcacacct	cccttggtcg	cctcagctcc	tgctctctca	actccagggc	70920
tcggtccctt	ccctgggttg	cctcttattc	cttttaaagc	agtgtttttc	aactagaagg	70980
gattgcaaat	ggcatttggt	agtgtttaga	gacagttttg	attgttatgg	ctgccagcat	71040
ctagtaaagg	ctaaacctac	agtgcacagg	accgcctcca	cagtggagag	acccaagtta	71100
gctatgtgaa	ggctgagaat	ccctgctttg	gagattaaaa	aagggaagctg	agggaaccac	71160
tcagttggaa	gcaccccttg	tggcatgcac	aaggccctgg	ttctgtccct	agctctgcac	71220
aaaaaataga	atacaaggaa	gagtaaacct	aatgagctgg	tcctcaccct	agtgtgccac	71280
tgaggtcact	tgaagggaag	tctagcccca	atttagtatt	ttttgtggct	gccatacctc	71340
cagccttgat	caaactctcat	ggtatacatt	ggtaagaaaa	agggtttgaa	acatagacct	71400

gatactcgga	catggaaaca	gtatgtttgg	tcagagagag	cgaaggacct	gatagacgag	71460
ggcaatatca	gagagagggc	atcagtcggg	ttagacacga	gcattccaca	gtgagcagct	71520
ctggataagc	ttttataaat	gctggttaag	gttttgaatt	tgcccaattt	tgtcaggatc	71580
ccagagtcta	tcacaaacat	acacagtttt	ctcaaatctg	ctttgcagta	tgcccgtgaa	71640
tgtctcttat	ctatactttc	agatggtaag	accctgaggg	cagaggaact	cagacccttt	71700
gtgccccctg	taagaccctg	ggggatgcag	tggaccgcag	tttgtgttct	ctgcacagaa	71760
aggagtccac	tttcgttgag	actaaggaag	ggaactgaca	agcttccctt	tctggcttca	71820
ggttggcagt	gcctgaagct	ctgagtgcca	tctggaaggg	agaaaggatc	ctgttggtgc	71880
actgtggccc	tgagagttag	gaggtccggg	agctccggga	aatgaggtcc	tgcttactgg	71940
actcgtcatt	cgtccctgtg	gggcccttgg	tgtagagcaa	tcatcctcac	cctcaagaa	72000
gagctctggt	gtgactgag	atgttctggt	ggcttgagc	tctcatcaga	gaggacggga	72060
ccttcccacc	tgacctgagc	ctagtgtctg	gcacagagag	cacttgaaaa	cagattgaga	72120
cactcacctg	ccatgctggc	tgtctgcttg	aagagctaac	tgccctctga	tggaaccccc	72180
atgcccagaa	aagactaaat	ccagtatcta	aaggttgctt	taaagggttg	tcactgcagc	72240
cgggcttggt	ggcacacgcc	tttaatccca	gcactcggga	ggcaggcgga	tttctgagtt	72300
caaggccagc	ctggtctaca	aagtgagttc	taggacagcc	agggctacag	agaaacccctg	72360
tcttgaaaaa	ccaaaaaaat	aaaaaataaa	aataagtaaa	aaataaataa	ataaaataat	72420
aaagggttgt	cactgatctg	caggcagctc	atgctagcct	aggcttttgg	ctcgatttca	72480
tctcactaaa	cgtatgaatct	gtttccctgg	aacattccta	tggtttctag	tagtaatgaa	72540
gtgctgtggt	ccactccagt	gagaacttca	attccttagtc	ttgtattata	attgaaaaat	72600
aatataatgc	aagaaatcag	tatgactgct	tacctcaaga	gacatacaat	tccacttaca	72660
atatcctgct	tctttaaatt	tttcattaa	actggtgata	tataatttgt	gaatggagaa	72720
ataaatacgt	cttactgttg	gcagtttctt	cctgggatgg	caactctgta	ttggttctct	72780
accagtgtcc	taattctttac	tcagtggctt	tcattgagtg	ttcttggcac	tcactgtcca	72840
agcactgatg	caaggcaacc	ctgtagcatg	acttcatagc	acaggccctc	ttgttagcac	72900
acctgaaagc	agaccactct	ggctgtttca	cttgacagca	gaatcttact	ctgtaagcca	72960
gtctagcctc	aaacaacatc	ctcctgcctc	agccttccaa	gttctaggtt	tataggaaaa	73020
ggccaccttg	cccagcttga	gactgcttct	tactgccatg	tctcttcagg	ctcacacatg	73080
aagtccaggg	cactccagga	ggagccgtga	gtctgtctgc	agggcactcc	agggggagcc	73140
atgagtctgt	ctgcagggca	ctccaggagg	agccgtgagt	ctgtctgcag	ggcactccag	73200
gggaagccgt	gagtctgtct	gcagggcact	ccagggggag	ccatgagtct	gtctgcaggg	73260
cactccaggg	ggagccatga	gtctgtctgc	aaggcattcc	aagagcagcc	atgggcgtca	73320
ctcattggta	gactgtgagg	ctacatctcc	agatgccccc	agtgtgttgg	ttgtgagcac	73380
tgtgtctcat	ggtttccaac	tgagacagag	ggaaggactt	tgcccccttc	cctaaggatg	73440
ggtagtaata	gtccagacca	caagggacag	atagctatgg	ggttttctga	ctcatcctta	73500
gtacattatt	gctgatgacc	agtttgtttg	gatgagttag	tgggaaagaa	gacccaagtc	73560
catacactct	gctttttaga	acttgctcat	cctagccatg	cccaaggagc	agccgttgac	73620
tgtcatggca	ttacagttag	gaaataaaca	gtcctgaagg	tgccctggcag	cagcttttca	73680
agaagctggt	gttaaaagac	agtattcaaa	catctgcgga	ctgggaactg	ggcagcattt	73740
gagtctcctg	ctgtctgtta	atttaccctg	acaaggaggt	gacttgaaag	gtttgttttg	73800
tttggggtag	agctttttca	ggaaaaaagt	ttagtcttac	agacaactct	atagtatttc	73860
tagtccaaac	tcatgccttg	tgttttattc	ctaaaagccc	tgtcacactt	tgtaaaatag	73920
gtgctcttcc	tcaaaggata	tatttaacgt	tttatatc	aggccttatt	ctgtgcattg	73980
aagctttttt	tagatgcttt	gtaagatggc	tcagtggtta	agagcatgta	ctgtcttct	74040
gggaagtcctg	ggtttgattc	tcagcagcta	acaccagctg	ttattccagt	tcctgggatc	74100
tgatgccctc	ttctggccta	tgtgagcact	gcattgtcgt	agtgcacaga	caaatgcagg	74160
caaagcactc	atacataaaa	ctaaattcaa	aaaactcttt	cattgtctca	tgtgacctag	74220
cttgagaata	cctgtgctta	tattataatc	tagtatgagc	cagccacggg	agcaacacac	74280
ctattatctc	agcactcaga	agattgagac	tagatgggtc	agagctagag	tctgggttac	74340
aaaacacctg	tctcaaaaagt	aaaaggcgtg	aaaaagtgtc	tcagcagcta	agagcacaca	74400
ctgcttctcc	agagggcctc	atttcagttc	ctaataccca	caccgagtga	ctcaaccacc	74460
tgtaaactcca	ggtccatgag	atccaacacc	tctgggtcgtc	tgcataagct	cctacactca	74520
attatacaga	gagagagaga	gagagagaga	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	74580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	74640
nnnnnnnnnn	ntctagagtg	tttcaggttt	ttttgttttt	tttttttttt	gagacaagg	74700
ctctctatta	tgtctgcttg	aactttctat	gtagaccagg	ctggactcaa	acttatagtg	74760
atccactact	tctgcctctc	agtactggta	ttgaaggcat	gtgtcaccac	acccacttac	74820
ttcaagatct	tagatttcca	aagaagccgt	agcctagaaa	aggttaataa	gtactgattt	74880
aaaacagaaa	gaaatcaggt	acacttagag	ctgtagaatg	tcagcatgtg	acatttgtga	74940
caagtgtgta	aaactttgct	cttaattcta	aaagagagaag	ctgtcaaaa	acttgactg	75000
gggctgtagc	caacttggtc	gagcccttgc	atgaagactgt	gtgtttactc	cccagcactg	75060
tggggtttga	attgatttga	accagtaga	ttcgtatatt	tgaatgttta	cctcatgggg	75120
aatgacatat	tacaaggtgt	ggccttgttg	gaggaattgt	caatttgggg	gtgagctttg	75180
aggtctctct	gctcaagctc	tgcccagggt	agaaaggag	cctcctcctg	gctgtctaca	75240

gaggacatag	tctcctggct	gccttcagat	caagatgtag	aactcctggc	tcctccagca	75300
ccaagtctgc	ctgcacaatg	ccatgcttcc	taccatgatg	ataatgaact	gaacctctga	75360
aactgtaagc	cagccccaat	taaagtgttg	tctttataag	agttgccttg	gtcatgggtg	75420
ctcttcataa	caataaaaagc	ctaactaaaa	cacatttcctg	ctgggagctg	gtgggtgcacg	75480
cctttaatcc	cagcacttgg	gaggcagagg	caggaggatt	tctgagttcg	aggccagcct	75540
gggtctacaga	gtgagttcca	gaacagccag	ggctacacag	agaaacctg	tctcaaaaaa	75600
aaaacaaaaa	caaacagca	aacaaatgcc	agcatttggg	aggtagagtt	aagaagattg	75660
ggagtacaaa	gtcgtctcag	ctagtatgtt	tgaggccagc	atggaccaca	tgagacgttc	75720
tcaaaacgaa	agaaacgaat	gaatagataa	acatttgagt	gtccagtttt	ttcctttctt	75780
tcttgctttg	tttttggcgg	tgctgaggat	taaaccagg	accttggtca	tactaggcaa	75840
gcattctcca	ctgaggaaca	ccctggcgag	tgcttagtct	gtctgtctgc	ctgacctgct	75900
gcctgcctgc	ctgcttggtta	tgtgtatgag	tggtaacctg	catgtctgtc	tgtataccac	75960
agacatgcct	ggatatctgca	gaggccagaa	gaggatgttg	gatcgcttg	aactgggatt	76020
acaaatgggt	gtaagctgcc	atgtaggtat	tcagaattga	acctgggtgct	ctgaaagagc	76080
agccagtgct	cttggtgttg	gttttattgg	gggcaggagg	tagttatttg	gttggttggt	76140
tggttggttg	gttggttggt	tttcttgaga	cagggtttct	ctatgtagcc	ttggctgtcc	76200
tggaacttgc	tctgtaggct	caaactcaga	gatctgcctg	ccctgcctc	ccgagtgcgt	76260
ggataaaagc	atgtgccacc	aactccagac	aagcagccag	tactcttaac	cactgagcca	76320
tcattccagc	ccttctttgt	gttttgagat	ggtcacaaag	tacaactcag	actgagctct	76380
tgatcaccct	ccctcagcct	cctgactgct	gggggttaca	ggtgtgtcac	tgctcctcaat	76440
tctgagtgtc	agatcttgaa	aacctattct	cgtgaccttg	atccttaaaa	caaaccttg	76500
gagaatgagt	tctgataact	atttctcact	cctcttcaag	aaaaggaaag	ccagagaaag	76560
gggggggggg	aagcccccaga	aacattgata	acttgcctca	agttacacag	caaaattcag	76620
acagcctgca	catctcagtg	gccatctgtg	ccatatccac	cctgaccttc	tctgacctcc	76680
ccacctccat	ccctacagac	cttgacgttg	agatcagagt	ccaagccgta	tcgtaagatg	76740
gccttaggat	ctgacatcat	ggggactctc	acggctcctg	cctcgcccaa	atgaaaatcc	76800
tgagggtgct	tctttctcga	gtcaaaactg	gttaccact	gccctggaag	gaaacagatg	76860
gagcatcctg	agccactgtc	cccagaaagg	ccacaggtcc	acgtgtgtcg	tccactggcc	76920
aaagcaacct	gagctgtcag	cagcaagaac	acaggagccg	ctgggtccca	gcatgtgtgg	76980
cacagaccat	aggctccatg	caccacgggt	tctggctatc	ctcctgtagt	aaactcagaa	77040
ataagtgggt	gttctctctc	tgacttggat	caccacgctc	cttctgttta	aagtggcctt	77100
taatattgctg	gtgtgtggta	cgtgcctgct	ctcctgtccc	ctggggactt	ggagttaggaa	77160
gccccagggc	tttctcttaa	gttaggatcc	actcttgcta	ctactccata	agatgggtcac	77220
aaagcaacgt	aaaatggaaa	ttaatcaaac	cattcctgcc	acaagaataa	aacagatctc	77280
aggggaggcc	tgtggaagg	tctcctgagg	ccttaccact	gtctagaagg	aagttgacag	77340
cagttcttga	gcaggggtgc	gactccagga	gttgggggct	gctctgagag	caggacagca	77400
tgtattgtag	agtgtctggg	agggagctgt	gttatcctta	ccgtgaagat	gaggacacgg	77460
gctcatgggg	gcagagccag	gattaaacct	ggtctgattc	aaaaagccag	agatctgtgc	77520
ccagccccac	gcagccattt	cactgggtcaa	ctaattcaga	aacacttggt	ctgatattgct	77580
catagcgtac	aagcactgtg	gccttcagat	ctcctctgag	cctggtaacct	gcattcaggt	77640
tcaccacat	caccacacac	acacacacac	acacacacac	acactcggcc	agagacaagt	77700
ggggaagccc	tcacccttga	agtaagccac	gccaaaggaga	aggatgctga	gggcaactgg	77760
catttccctc	gtggaccggg	caatcttccc	tttcatctgg	gcctgcaccc	agttgttaat	77820
ctcctgaagg	tctactcgag	ggttgcccgt	gaggatccgg	ggcctgggtc	cataggactt	77880
ctccagaggg	gcaacaaagc	tggaatttgc	tcgaagttct	tgagaggaaa	cagatcaaaag	77940
atgagagctg	aatcagcacc	ctcactttga	aagcatgcc	gacccagct	tcctgctcag	78000
catcttccct	tgacttgcgt	gggcactctgc	cggcttgccc	agaccctggc	tagggaacag	78060
tggaattccac	cgtttgcat	ccccgtccca	ggcctcctg	ctgtctccca	gagcccaact	78120
cctctttctg	ttcctctgtg	gtctcactgg	ctcttctcctg	cccaccagtg	ccaggcctcg	78180
cctgagcaca	cacagcctat	tgtttagaca	tcattggaagc	atacagacaa	cccaggccaa	78240
tgaagcaact	tcacgccagg	cataatgggg	cgtgcctgcc	cttcagaagc	agaggcagct	78300
ttatgagttg	ggggaccagc	tgagactcta	tagactttga	gagggtgggt	gggggtgggg	78360
ctactgactc	ctctcaaaca	caattctgga	agcactcttg	aggttcttct	caggggcagt	78420
aacagaggca	aggagctcct	tgtaggtgct	gtggatgtca	gggttggtga	tcaggctcgta	78480
gtagagagcc	cggtgaatga	cagactctgt	tcgatgttca	gctcctgcc	gagagaaaag	78540
gatgccaaagc	ttcataactg	ccggtgaggc	ccgcatacag	atagggacgt	tagacatcaa	78600
tccttttgct	ctctgagagc	ccgaggaggc	cgatattgca	gatgttttag	ctggacaaga	78660
tcttcagggc	gtggaaagaa	ataatgaccg	ccttgctagg	aagagctcta	agacagggca	78720
aggttatcag	agctacagag	agaagagtgg	gatgtggtcc	tgaagttctc	ccatcgtaac	78780
ctaccctgtt	ctgaggagga	gccagctctg	ctcacggcag	ctgtaccctc	agaacctggt	78840
taaatgacta	aaacacgata	ggaggccact	taaggaaacca	aggtcgagt	ccacttaca	78900
agtggtaggg	attgtgtgtg	tggccccac	cgcccccttc	ctgttctctc	gacggcgga	78960
gcatggaaac	tctgagtggtg	ggaaattcag	gtccacctgc	agccttcttc	agttgacact	79020
caccagaga	aagggcagag	agggccgtgg	ccacgctgag	tggagacagc	aggacgttgc	79080

ccgttgggct	ggcactggat	ctcagggcgt	acagatcgta	gccgaagtgt	gagacagctg	79140
ctgccagctt	gttcacaggg	accttgaaga	aggggtcctc	ctcctccacg	ggctcgcccg	79200
tgctgtccgg	gactggggag	ccctgggtta	gaatacaagg	accagtaggg	aggcacagtg	79260
agtacatcac	ctcctgggtg	ggttgggtcct	ctagtccctg	gggccatgag	tctgaggtca	79320
gaatgagtgt	gtgtctctctg	actccacaac	ctgtgtgctg	ggaggtgggg	agtgggaagg	79380
gcaacacaaa	agggccttgcc	agacctgaac	tgtgggtctga	gaacctgaag	cctggcccac	79440
tttaaaataa	aacttgtagg	gctggggaga	tagcacagta	gataaagtac	cagcatgcaa	79500
gttcaaggac	ctgggttcag	tccccagagc	tgggcacggg	ggtgcatgct	tataatccca	79560
acactgggga	ggcagagatg	ggcagggtcct	ggggctcatt	ggccaatcag	cctgaactaa	79620
tcagcgtatt	ccatctcagt	gaggggtcct	gtttcagagg	gcctgaggaa	tgactctggg	79680
ttgactacta	gcctactctg	tgtctgtttc	tgtctgtctg	tctgtctgtc	tgtctgtctc	79740
cacccctctc	gtgtcccttc	cctctgcagg	gaacttcctc	accaccacca	acccccaag	79800
aaaccacccc	tcagaccagt	cttccctatt	cagcttgctg	gctggtccta	gtctgcctag	79860
gttctgtgtg	gacgcctcc	ctgtctttcc	tgacaagcca	tccctctga	ctagaccoga	79920
gaggaatttg	tcgttttctg	acctgttttc	agtgtcagcc	tcttcttctat	gagactttct	79980
gcttttttgt	tttgtttccca	gggctttgga	tcaaagctgg	gctcttacat	acgttaggca	80040
aatgcttgcc	caccagctg	tacctccgt	ccctgttgct	tttcgggttg	gaggactttt	80100
tttttagttt	tctgtttggt	ttttggtttt	gattttttgt	tgtttggttg	tttgttttga	80160
gacaggattt	cgctatgtga	ctctagctgt	cctgggactc	actatgtaga	ccaggctggc	80220
cttagattca	gagatccacc	tgccctctgcc	tcctaagtgc	tgggattttt	agattttaat	80280
ctgtacctac	caacctcaaa	ggaagtgtcc	atggatagag	ttcagtacta	catcatgtgt	80340
gtaacatgtg	tgagggcctg	ggcttcaccc	ccaacagaga	agggggagtt	agtaggtgag	80400
gaataagtag	ctggctagt	gcaagacagt	attgtctaag	gtcactaagc	cttaagccac	80460
acttaaagcg	cacaatccag	gtctaataatg	cccactctgcc	ttgtccttgt	gtgacatgac	80520
cccaccccta	cttctctcgt	atagtggcag	ctcctctgga	tcctgaaagg	agaggggaaga	80580
tattcttgtc	tcgatgttaa	agtaaccaag	gcctagaaga	gtgaaggcca	aagccacccc	80640
tggtatccagg	gctgcctccc	tgcaactgtct	cttctgtgtg	cccacctaacc	caccctaactg	80700
acctcagagc	tgctggggac	gttctggctg	ctgcccgtgc	cgagcagggc	tccagtccag	80760
aggagttagc	ccagggcctg	catcccggaa	ctacaagaga	aacaagagag	caaacgactc	80820
ccctcaccca	caccctcccc	tgccaactgca	cattgcacac	tgcacaggga	caagagtcat	80880
gcagagttag	cccttccccc	ccctccagct	ctcagcccca	agtggaccct	tgacttgagg	80940
tcttccgtcc	ctgacctgcc	cctgcacttc	tccttgagct	gtgcccccat	gttggttctt	81000
atcaggagac	ccaccttccc	atctaagctc	cagcacaggg	aagaccagc	agcaggctct	81060
tcaggcccca	agacaatgct	ctagcacaaa	cacacaccaa	ggcttttccg	tggaaggcaca	81120
cggccagctc	ctttggtagg	atttggaacc	ctgtctcagg	atgggagcag	agcccaggctc	81180
atagacttac	agaacatctg	gtctggctcct	gctatccacc	aatagttctc	tgaccaaagc	81240
ctatgttaaa	gacacacaca	cacttttcttc	taggttaggtt	cttgtgtatg	tagcccaggc	81300
tagccttgaa	gttgcaagct	ccctacttca	tttgccctcct	gagtactaca	atgccagggtg	81360
tgagccatca	tgccctgactt	gactcccttc	ttctatttca	agagcaattc	ttagttaaga	81420
gggtcatgac	cagggccaca	ctgcccnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	81480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	81540
nnnnntctta	tttttagctt	gtttgttttt	cttacttgag	acctggggag	gggaggtgta	81600
tgtgtcttcc	atttgcttct	tctacctaat	aaagtttctt	ggtgatgttt	gggggggggg	81660
gagggggtag	gactcaagaa	gggattctcc	cataagctgt	tccgtttggg	tagtactatg	81720
taaggaaagt	acaggtgggc	agagctcggc	tctgcctgac	tggcgtgtct	tgaggtaaag	81780
gtgagatgtg	gcaagatttg	ggccctcagg	agttggctct	gttggccctg	taccttctgg	81840
tctgtgggta	aggatgacca	gtaggtgaga	gatgagggaa	ccagaacaga	aggtgaaagt	81900
tagtggggcg	gagccccaga	ctagtcagg	gggggtaaac	tagatgactt	tctggaaccc	81960
caaggggctc	ggagactagt	ggtgttgagg	aagacctcta	atgtgttgta	aggcctctga	82020
actcagtagc	cgaacttgat	gccagaaagc	cccaaactgc	taaacccaag	caggagcggg	82080
acgccatccg	ttccatggct	tcacccgagg	tggcccccag	gctgcgccaa	tcaatgagca	82140
gccgagagat	aggggctggt	acaagccagg	aaaagttaca	gcacgctgga	aagataatac	82200
aggccaggaa	gccccaggca	cagcagggtg	gaaaagctag	atcccgaattc	tgccggaggg	82260
gggccccttc	gaggtcccgg	gcacccgggtg	ccaggatcag	agaaactgac	tgaacactag	82320
ctgacctgcc	cagaccatgg	catcctgggg	actccttgtg	gctggcgctt	ccttcacggc	82380
gtttcgggga	ctgcactggg	ggctgcagct	gctgccacc	ccgaaatctg	ttcgggaccg	82440
ctggatgtgg	cggaacattt	tcgtttcgt	gatacacagc	ctactctctg	gagttagggc	82500
gctggtcggg	tgcggaactt	ggggactgac	aaagcactga	ggggcggggg	tggaaaagag	82560
ggcctggaag	actgaagttg	gaaccttttg	gaatggaaact	ggtttgggtt	gtgtaggggt	82620
gggagtaccc	agtgggagaa	tggatctagg	tctgggagaa	attgacctta	gctctttgtc	82680
ttctccaggg	tgtggcagtt	tcctcaaatg	gtcaccgacc	caattaatga	tcaccaccg	82740
tggtccggg	tcctagtagc	agtgtcagtg	ggtgagtgta	cagaaaaggc	tgaatcggga	82800
aaggccttgt	tggaccggga	attctaggtt	cctcccccat	ctttggaatg	gagcagatgt	82860
tgctggaggt	ttgctgtgag	gaattaagga	cctgagaaaa	gtgggacttg	agatatctag	82920

gctgtgcatc	agctctgagc	gaggagcctc	atagttcttct	cgggtgcctt	caggttatatt	82980
cgctgcagat	ggagttgata	tgctgtggaa	ccagacattg	gcccaggcct	gggaccttct	83040
ctgtcaccat	ttggcggtaa	gactctgaag	ggagaggcca	ggtagtaagg	gagcatgtcc	83100
aactcaaggg	cccaacctct	ctcttcagtg	ttctgtcctc	tgacttttcc	acaaagcccc	83160
ctgaaaacct	atcctctcag	acttgattg	agttggagg	aggttttgac	tggttagcca	83220
ctcctgggca	ctgcccagg	agtttggttc	tccccacaaa	cctccagctg	atcataaaaa	83280
aaaaaaaaaa	aaagccagga	atgaaagcta	gggtatgcta	tgcaaatagt	gtggcttggg	83340
gtaagagaac	ctctgggtcca	gggctgtcca	tgccccctag	ataagggtca	gcagaaaagg	83400
caggattgga	ggcagtccta	aaaaatgctt	gggtaataata	aagtgaataa	ataaaaaata	83460
aataaatact	aatttttaaa	aagctgatac	ctggaaggat	gaggcagaga	gtagaaaaaa	83520
catgcgtggg	tgctccctagg	ataaggagct	gggacttgtt	gggcacagg	catgcaaac	83580
ctgaaccttg	aaccttgccct	gcaggtagtg	agctgcctca	gcaccgctgt	tgtgtctggc	83640
cactatgtgg	gcttctctat	ggtatccctg	cttctggagc	tgaactccat	ctgtttgcat	83700
ctacggaagc	tactgtctgt	ctcccataag	gccccatcct	tggccttcag	agtaagcagt	83760
tgggccagcc	tggccacctc	ggtcctcttc	cgccttctgc	ctctgggatg	gatgagtctg	83820
tgtttgtccc	ggcagcacta	ccagctgtct	cttgctctgg	ttctgctttg	tgtggctggg	83880
ctgggtcaccg	tgggcagcat	aagcatctcc	acagggatcc	gaattctgac	caaggatatc	83940
ttgcagtctc	tgccctaccc	gtttatcctc	atgcaaaagg	aaaccaagac	acgtgagcct	84000
gttgccagga	acacttccac	tctcagctctg	aaaggtgtgg	aagttttctc	ttctgtcagc	84060
ccccagggag	gtggggctgg	gaagaggaga	tggtagccca	ctgcatagtc	tactatgtag	84120
caaggactag	actgtatcat	cagagagaga	gagagagaga	gagagagaga	gagagagaga	84180
gagagagaga	gaacattgta	tgagatctcc	attacagtca	ggaaatcagg	agatctaaat	84240
aaactttaaaa	gtcccacagt	ctttacatat	tcttaaaatt	tcaatctctt	taaaatatcc	84300
atctctttta	aaattcaaag	tctttttaca	attaaaagtc	tcaactgtgg	gctccactaa	84360
aacagtttct	tccttcaaga	gggaaaatat	cagggcacag	tcacaatcaa	aagcaaaagt	84420
caatctccaa	ccgtccaatg	tctgggatac	aactcacgat	cttctgggct	cctccaaggg	84480
cttgggtcac	ttctccagcc	aggccctttg	tagcacacgc	gtcatcctct	aggctccaga	84540
tacctgtact	ccactgtctg	tgctgtctct	ggtgtgtcatc	tcattggtact	ggcatctcca	84600
aaacgctgca	tgacccttcc	agtccctggc	cttcaagaga	gaagactaga	gcctggcaaa	84660
gtggcacatg	ctgataatgc	tagcacttgg	gaatgacaag	cagaaggatc	agaagttcaa	84720
ggccagcctg	ggctacaaga	gactctgttt	caacaacaaa	acaaacaaac	aaaccaaaaga	84780
agagaaaaga	aaaactggac	atgacagccg	gaacattatc	tgacattcat	aaggtcctga	84840
gttcaatgcc	aagttggcag	tgccctagtt	gataagggtc	tagccactct	ggtaatacca	84900
tggctgactg	aacaccttac	ccagcaactt	gctgatatag	tctgccttcc	agcaaaagg	84960
aggagcttcg	ctagggagag	aacattgaac	cctattgtat	atgaataaat	tgctgtgcaa	85020
atgatttcat	cagtctcttg	tgaatgtgat	tgctttgagt	catttttctt	ggctccagtg	85080
ttatcctggg	ctgcagtgtg	gtgtggagtt	gtggaagctt	tgagtgggga	gggtttcctg	85140
ttaagggttc	tctggtctct	ttctttcctc	cgggtttttg	ttttgtttgc	ctggtggggg	85200
tctctggtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	tagaagttgg	85260
cgggggggtg	agggggctgg	agagatggct	cagcgggttaa	gagcgccaac	tgctcttcca	85320
aaggtcctga	gttcaaatcc	caacaaccac	atggtggctc	acaaccatcc	gtaacaaaaa	85380
aatctgatgc	cctcttctgg	agtgtctgaa	aacagctaca	gtgtgcttac	atataataaa	85440
taaataaata	ttaaaaaaa	aagaagttgg	catggatgat	gtagtgaaga	ctggcattag	85500
atatctctgg	atccccctgc	ctctacctct	tagacactgt	gagtatggaa	gtgtaccacc	85560
gcaccaggcc	aggctagaac	attctctgat	ctacaaaatac	ctagagtatt	attcctctat	85620
gatcagaaaa	cagaccaggg	gggccacaga	aatgtcttag	taggtaaaaa	cacttgcttt	85680
caggcctgat	aacctgcggg	tttttggttg	ttctgggggg	cgggagaggc	tggctggctg	85740
gctggcctgg	aattcacaga	gatccacctg	cctctgcctc	ctgagtgtca	ggtaccagga	85800
tcacagggtg	gtgccaccac	acttggccta	actgcttag	tttgagcatc	agtactcaca	85860
tgttactgag	gatagaatag	actctcacca	gctctctctga	cttccacatg	tgccctgcag	85920
catgggctct	ccttcccaa	aggaaaaata	aatgtaagaa	ttaaaaaaa	aaaaaaaaag	85980
caaacccagg	tcttgtgtga	tggctcagca	tcaaagctac	ctcccgccac	agctgaccac	86040
ctggtgataa	cttatagcct	tgttatgtct	tcctttgacc	tccacgggca	tgctgtacac	86100
gtatgtgtgc	ccacacaaac	acacaatcaa	gaaataaatg	cagccaggcg	aggtggcaca	86160
cccctttaat	cccagcactt	gggaggcaga	ggcagggtga	attctgagtt	cgaggccaac	86220
ctggtctaca	aagtgagttc	caggacagcc	agagctacac	agagaaaccc	tgtttcgaaa	86280
taacacaaaa	aaaaccactt	taaatattat	ttttattttg	ttttgtttat	cctggaactt	86340
ggtctgcaga	ccaggctggc	cttgaactca	cagagatcca	actgcttctg	cttcccaagc	86400
acattaaagg	atgtcccacc	actgcctggc	taaagattta	tttttctctt	ctttttgttt	86460
tgttttgttt	tgttttttct	aaaaaatttt	tttaaaaaga	accatccctc	ctagactca	86520
ggagactctg	aagtcagggc	cagccagggtc	tactagtgga	gctctagggc	agccagggct	86580
ccacaaagaa	acctctctct	aacaaacaaa	caaaagagaa	cagacccaac	cagacctgag	86640
gacacacact	tgtaatctaa	gcccttgaga	ggctgagaag	ttcaaggcta	gccacaagtg	86700
tgtggtgcat	tcaagagcag	cctgggtggg	ctacagaaaa	agaaaagagg	agagagagaa	86760

tggttaaatga	agatgactct	ggaaaagtga	aactcaagag	aaagcccctc	agattttgctt	86820
aagacgagtt	gaggggtggag	aaccgcacaa	gcgagcagag	cagacagaga	ctgccacaa	86880
agttcaatcg	gttcagggtac	attacttcca	aaacgccatt	gccacatcag	gatgcttcaa	86940
tcagccaaac	caacgcagcg	actattgact	tctgcatttc	agagacttcc	gtctctgtcc	87000
agggcaatgt	cacttttagct	ttcctttgca	gaaaggaaaa	gtccctgcct	ctgatgtggt	87060
agatcctcac	acaccttctg	ccagatccag	acactggtat	gactcagcct	cggggagctc	87120
tatctacaga	gataagggtg	caaggcgtgt	gtgttttaag	tatgtgttta	aaagtacaaa	87180
gtgagagtcc	ctggaaaggg	ctccctgccc	tcaccatcac	cgaaagcaca	aaccttaggg	87240
taatatctga	cattcctgga	aatgtatgta	tgtattcatt	atgtagccct	gactgtcctg	87300
gaatggggta	taaaccagga	tggcttcaca	tctcagagac	ccatttgccct	ctgcctccca	87360
agaactaaga	ttagaggcat	gcactacat	acttggctca	tgatttactt	aacttttatt	87420
tatgttcacg	aatgttagcc	tgcatgtatg	tgtgtgcacc	atgtgcatgc	ctgggtcccc	87480
agaggccaga	agaagggtgt	ggttggattt	cctggagatg	aagtcccaaa	caactgtaag	87540
cagtccaatg	tgtgtgctgg	agatgaaact	tggttcatcc	acaagagcag	tatgtgtctt	87600
taactgtgga	ggcatatctc	cagcctcaga	tttcccagtt	aatgtttgct	ttcgcaccca	87660
ggccccatctg	cgcattgcgt	ggagacctcc	tttaccgcct	tgagcctcat	tggccaattg	87720
tggctgggag	acttgcagat	cccaagtgtg	acaagagaag	aataaaactg	tgtgctatga	87780
actcacctct	tctctgtagc	cattggctga	gcatactttg	cctcaaccta	ccgcccttcc	87840
ttccccctaat	cctaaatctt	tgcctctctc	aaatgtgctc	ctcccccgca	gtaatccagt	87900
ggctcgctggg	gctctagaga	gatggggggg	gggggagcaa	cgggtacagc	ttaaggcagc	87960
tgacgacagaa	cttttttgcct	gtatattgag	tcttaaaaat	tcatataaac	tttgtgttct	88020
gtttctaaat	ataaccccat	ctgttttcaac	acaaaaatgca	acaaaaaat	gtttcaaat	88080
gctatttggga	ataatttaaaa	aatttcaata	cttgattttaa	aaatgcttta	actttttaaa	88140
taaatttttaa	atgttattat	ttttaaaaag	ttacaagttt	aaaaaaaaga	aagatagaaa	88200
tcacataatg	aaattaacca	tacgcaagtg	aggctcgggtg	cactggtaca	cagttacagt	88260
agccatttgg	agtggaggcc	atggcgcttc	acattgaatt	ttatactttc	tttatagaat	88320
attttttatg	cacctatcta	ctactgataa	caaaacaccc	atgagagagt	tagaattaga	88380
catcaattag	ctttgatcct	ctgtcataac	togtgtccac	tcctgcctt	agtcctacct	88440
catccctgtc	ctcttttcta	catcttatac	tgaatccaca	cactcagttg	tttacacaaa	88500
cacatacatc	actgtccann	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88620
tgccctggctc	tgttttgctg	ttgttgtttt	tgtttgtgtg	tggttttaga	tttggttttg	88680
tttggtttga	ttttgttttt	tttttagagaa	tcttactatg	tagctcaggc	tgtccttgaa	88740
ctcacagaga	tctcttgtct	ctgccttcca	agtgctgaga	ttaaagggtat	acaccacctt	88800
acctggcccc	tttcatctat	ctatctatct	atctatctat	ctatctatct	atctatctat	88860
tatctatcta	tctaaaattt	atctgtgtgt	gtctgtgtgt	acattcccca	gagcctgtgt	88920
ggtagtcaat	aagtaaccct	cagaagttgg	ctttctctaa	tccttggatc	aaacttgaat	88980
tgtttaggctt	ggtagcaagc	atgtttaccc	actgagccat	ttatgacccc	atggccagc	89040
actttccatg	ggttctgggg	acacaaatgt	gtactttgat	gtttacagga	caagcgctta	89100
accaaccaag	tcattttccc	agccccatcc	tgactcccat	taagtgttct	ttcccccaac	89160
ccaggaccaa	atctagagga	gtgtccatgc	ccaacaaaca	ctctgccaa	cctctcccct	89220
tactgtctct	ctcccttccc	ttccttcatt	tcttctgtcc	ttcttttctt	tctttttgaa	89280
acaggtcttt	tctctgcatc	ccaagctagc	cttgaacttg	tgatgtagct	caggctggct	89340
ttgaactcac	agctgtcctc	ctacttcagc	ttcccaaaaca	ctgggattat	agacctatgc	89400
taccacacct	ggctcatttt	tcaataaat	aaaaagaaaa	tcaaaaagtt	cctagaacag	89460
tcacaggatt	cacaaaaact	ttggaaggag	actaaaaatg	gattttttaa	aaatgcttga	89520
agcacaaaaga	gttgttgaaa	gaagagagaa	gaggaaaaagt	tagcttagta	ggtagaagtc	89580
aatcaagcct	cacaccctga	gttcaattcc	tgaccctatg	gtagaaggag	aagagcaatg	89640
ccggaaacat	tatcctctga	cctccagacc	cgctgtggca	cgtgcatgca	cacacacaag	89700
caggagccct	tggagggaag	tcctagaaat	gaactttact	gaagcagggtc	tgccaggccc	89760
tgtgctcagc	catttttattt	ttccttttgt	taccgcacac	gcttccattc	tcaaagttgt	89820
gagtctgaga	ggaagtactc	actgtgtccc	cagttagctt	ctgtcttacc	ctgggtcact	89880
tagatggggt	cacttagtgg	tagccttggg	gtggagaaag	agaacacagg	tccgagtagc	89940
cagtagacct	gagtctttat	atctgcaaag	gggtgtgggg	cataatcaaa	tctccccccc	90000
tccccggggt	cctgatacca	ggttgtatta	agtgtatgtg	catggtcctt	ccaagtcttg	90060
acacatcatc	caactccaag	tggctttctc	atttttccct	gccagtagcc	tcttggtgag	90120
gaaatggctg	aggaaaacag	agttgcagaa	agacagggcc	atggcctggc	tgcaggcttt	90180
ctctgagtct	gaagagggtc	agcgactctg	agaaatgaag	ctatttctga	gtgagagggg	90240
ccaaagaagg	aacacggcag	agggagagcc	cccaggagga	tggagacaga	agccggagag	90300
ggaccctgtg	cgaggctgga	ggggaggaag	aggggggag	agtgagaccc	actgtcatct	90360
gttgggcaga	gaggggctac	attcatctgc	agtatggtgt	agaggggaca	gagagtgatg	90420
gtaacaggaa	aaatttgggg	ttgagggggg	cagcctgtag	ggctggggcc	cagcagtgta	90480
cagctaggta	gagtacacag	taactcccag	aattctctgg	ctccactaaa	tccctgttcc	90540
gctccgtgca	gagtaaaacc	cacacagggt	ggatttcagt	ctcctttgca	ccccctcca	90600

ccccccctcc	acccccagct	ctgggtcacag	ccagtcagag	ttgggggtgg	ggcagatcctt	90660
gtaaaagagg	ctgggtgagga	catcggaag	tctgtaccct	ccactagcaa	agtgcagagac	90720
gctccgtgac	acttttaaatg	cctcagataa	aacagtgaga	gactctcctg	gtggcaggca	90780
agatgatggg	tcaggagacct	cagcgccctct	gaggctcaga	caccaggata	aagaataaaa	90840
acaccacgga	gacccctgtg	acccctcctg	gcagagggag	aatgccaatg	tggcccagct	90900
agctgctgga	gcgcaagcct	caaggcctgt	gctagttaatg	agtctactgc	tgctgccttg	90960
tcccaataaaa	cccctttccg	cccaggatta	gtggacacgc	cttgctcaag	ccgagtcctt	91020
gatctgccac	cttatcacac	acatacaaaa	atcccttgag	gatggttacc	atcctgggac	91080
aaagcctcat	tctctgctta	acccaagtga	cacctatatg	gcagatccct	gtgtccttct	91140
cctgatgata	acaacccttg	caatccaata	gaggggaaact	cgggtttctg	tcagcttccct	91200
ttatgctgat	agaaatgtac	tctgcatgtg	gggagcctgc	cttgctcacc	ctgagacccc	91260
atggggctgg	ctggggcttt	gcacatcatt	gggactcaga	gatgttgact	acatgaacgt	91320
cccacacttg	gttgacacaag	gcagaatgac	aggatgttat	gcctgggtgtg	tgagtgtgtg	91380
tgtgtctctg	tgtgtgtaaa	acgcctctct	ctggagccct	cctgtctgtc	tgccctcttg	91440
tcaatggctg	cacaattgtc	ctttctcttt	ccaaggacct	ctgtatgggt	gtgtccttca	91500
ttcagtgcc	ttcctctgtg	ggtttgtcct	gctagccccc	tgtcactgag	aaagtcttct	91560
gtctgtcctt	gggttgtctg	gctagaacac	agacatcatt	gtcttttttt	tttttttttt	91620
ttttttttta	gattttattt	atatgtaagt	acactgtagc	tgtcttcaga	cactccagaa	91680
gagggagtca	gatctcgcta	ggatggttgt	gagccaccat	gtggttgctg	ggatttgaac	91740
tccagacctt	cggaagagca	gtcgggtgct	cttactcact	gagccatctc	accagccccg	91800
acatcattgt	cttgcccacg	actgctctcc	agaatggtgg	gcaggaggat	gtgaccccc	91860
acccccaggc	accgggacac	aacatcttct	acacgtgtag	gtcttgtgca	ctggctttgc	91920
tttcttcttc	aagcagggtc	tcccaggaaa	tggcacttac	agagattgaa	gagtttaata	91980
catgtctcgc	tgccctctct	ttcgggaacc	ccccagaggg	agcagcagaa	accagggtcg	92040
gcaggggctc	taagctgcct	gggcaaagga	gcagggggta	gcatggagcc	ttagccaatt	92100
tggaaagcac	tgtgacccaa	gcacattttg	cagcagtaat	gtcaaattct	gccgttcagg	92160
catgccattg	atgtgcacgc	tgccacacag	aaaccagtga	cacaaaggca	cagccttctc	92220
caccctcctg	gtgcttagga	actaacggct	ctaattgaga	atgagagctg	aaaggagaga	92280
gacgggggag	ggccacagca	gcgcaggctg	gcactgcgtg	ttggaggagg	ctgacccact	92340
tctcgtagag	gtaaggggac	cactgaaatg	tcacttaaat	tagccaccac	tcccaacact	92400
agatctcctt	tgtccccata	cctcagcccc	acgcttcttt	ctttttttct	tctttttctt	92460
ctcctctggg	gcagcctcaa	gcccagcacc	cacttttttag	agctgtaaac	caccctgggtc	92520
ctagaagccc	tcttacgtta	ggggatgaca	ggaggtagag	atcaggaagg	agggagggag	92580
gggaggagga	aaggaaaagg	gaggggagag	agggaaaagag	atcgagagag	catgcatcca	92640
tcacaaagag	ccctcttttc	tggctttttg	actgcactgt	gagttattta	gccaacaata	92700
gatgtttatg	tattttttta	gaaccctgat	ttattaacag	cctgaaagga	gagagacgga	92760
gatttatata	ggaagtgcag	tgagttaagg	ggggcaatta	agagagcaga	aagagatacg	92820
gaacacagac	ttgtaaaggg	ttttgttaaca	tccaatcaaa	ggtgcttcag	gtattttcca	92880
aggaagcaga	aggtaaaaaa	aaaaaaaaat	tgtcccatga	gaagctgaca	ctggatggag	92940
caatggccca	ggcggaactc	ctgcttgaaa	gaagggtgaga	agggagggac	acagaccagg	93000
atccgatgag	ccagagtgtg	gccatagctg	ggtcatgagg	cccagggttg	gaaggacccc	93060
actaaagtgt	gcactggcct	ttccttgaca	aaggatgcac	ctatagctag	gcgtgggtggc	93120
aagtggttgt	tattctagta	cttaggaggc	tgaggcagga	ggatccaccat	gagtgtatgc	93180
ccagcctgga	ctgcatagca	acaccaggtt	tcaaaataac	aacaaaagga	agtgggggtg	93240
gggaggggca	cattttgaat	gtaaatataat	aaaaaatctt	ttttaaaaaa	agaaaggggc	93300
tagtgagtta	gttcagcgtg	taagagcgct	gactgctctt	ccgaagggtc	tgagttcaaa	93360
tcccaacaac	cacatggtgg	ctcacaacca	tccataagga	gatctacgcc	ctcttctggt	93420
gtgtttaaag	tcagctacaa	tgtacttaca	tataataata	aataaatctt	ggagtgaagg	93480
ggccagagca	agtagaggtc	ctgagtttaa	ttcccagcaa	ccacatgatg	gctcacaacc	93540
atctgtacaa	ttacagtgca	ctcatatata	taaaaataat	aaataaatct	ttaaaaaaag	93600
aagaaagagg	gtggggcagg	ggagggaaga	agaagaaaag	taagaagcta	aataaaaagc	93660
acagagatga	gcttcatgtg	gaaacacagg	cctgtagtcc	tggcactcag	gtgggggttg	93720
ggggggctac	agtgaagta	tcagtgttcc	aaggtaact	tgggtgagac	cttgtctcaa	93780
aaaatacata	ngcnaaaaaa	aaaaaaaaaa	acatagccag	gcatgatggt	atacatttat	93840
agtcccagca	cttagaggac	tgaggcaggg	cagaaagaaa	aggaattcaa	gatcaggctg	93900
agctgtatgc	agtcctgatc	ctatcccctc	ccccccccc	ccagagacag	acagacagac	93960
agacagacag	agagaaaacac	aaagaaaagg	gccttcagat	ggctcagcaa	ttaaaggcgc	94020
ttgctattca	gaccccatga	cctgagctca	aagcctggga	cccaaggtag	aaggcaagag	94080
ccaactccac	agagctgttc	tatgatctct	atatgaatgc	tggggcatgt	gcctacacta	94140
tgtttgtgcac	acatgcacag	attagaaaaa	gaggaggaag	aaaaacataa	gattgtttca	94200
agaaaagaaa	ggctggcttc	ttccacgtca	gtgtgagagg	agggtctggc	ccctttgtag	94260
ccaggtcctt	cccagtcag	tgggggctga	actgaggcag	cggaggaggc	aataacggag	94320
ctttcccaac	gcagtgtcca	gcaaactcaa	ctctacagcc	tgtcctgatc	cacagagaag	94380
ccttcctggc	tcctccacca	atgcgggggc	attggctccc	aggctcctgg	gccccccccc	94440

acacctgtgg	agtgttaggt	gatttgctaa	tgttgggcaa	catttgccca	cgtgggggttc	94500
ttggctcttt	ggtaatagac	atgcctagca	ggaggggcga	gcttggaggg	gggagtcctg	94560
gggttgcccc	tggctccctg	cagctggggg	gtctggccag	ctgaagaagg	agccatggca	94620
cgcaaattgg	agagcatgga	acagaggctg	tggatgctaa	gcaatatggg	aggcagtcct	94680
agcttggaag	cagcaggtgt	ctgggaacgg	gcctgtggcc	caggcagatt	tccagtggagc	94740
actccagttt	tttggcaca	ggaacaagct	ggctgagccc	aagaggcaag	tggtgataat	94800
gaaacccgca	gttgagggaac	agcgggtaag	ggtgccatgg	gagcccatgt	gctcatgaag	94860
aggctggggg	gtgaagaaga	gcccattgca	ggaagccaca	catcccccct	agttccaggc	94920
agaggcagag	tccctgagtg	gggctccctg	ggtctcccct	tacctaacca	gtctcccggc	94980
accccagcaa	acaaaatccc	atccataatt	tgagggtttat	agagacctca	aaggctgagc	95040
tactgtgtgc	cactaaccat	cagcctaacc	ctccccact	gtcttctcta	gctgcccctc	95100
tttcttctga	gactgtgata	gtggcgggga	cgggttggga	gtgtgtgtga	agccctctcc	95160
gactctccaa	ccccagctga	gccccttggt	ctgcagctca	gtaacacagt	aacacaggct	95220
cagttctaca	ctgggttgaga	acactcacgg	ctctctcagc	tccttagaga	gcctgttttc	95280
tcatttttct	gtcccccagg	cctagacaat	ggctgggtcca	tttgtaagct	tatctgagga	95340
tgccaggggg	caccccatgt	ctccactagg	ctggcaatgt	tctctgtcac	tgtagtacag	95400
aagactgcct	gggtgggagg	gagataagga	aagggtagg	ctccccggg	gttccacac	95460
agtgtctgag	ggaaaatggc	agaatgggct	gggaggtaac	tctgttgcta	gagtacttgc	95520
ctagcatgtg	caaggaagg	cctgggttcc	atccccagca	ctacagaacc	caggcgtggt	95580
ggttcatgct	ggtattctca	acattcagga	ggtacagtca	ggaagagcag	aagttcaagg	95640
ccatcctcag	ctacatagct	agcttgagac	cagcctgggc	tatgtgagac	tttgtctcca	95700
acaaacaaca	acaaagcagc	agaaggccaa	ctggcaagag	gagtattacg	taaagtaaat	95760
ccatctcaaa	aagcaagtag	catgtatctt	cttctatttt	tttttacatt	ctataaaggc	95820
ctatcaagtc	atgtatatat	gcatgtatgt	ttgtatgata	tgaagtagg	gggctggata	95880
gatggctcag	cagttgagag	cacttgatg	ctctttcaaa	gaacctgggt	tcaattccta	95940
gcacccacat	ggcagctcac	aactgtctgt	aattccagtc	tcaggggatc	tggcacccctc	96000
acacagatat	ccacgcacat	aaacaccaat	gcacataaaa	tataataatt	ttaaaaaag	96060
aaattggaag	taaaactctc	taaggagaca	aaagggactg	aggggaagtg	ggaggggcat	96120
gaagggggag	ggcataggtg	tgtggtgtgt	ttacacatga	gaatacactt	ctataaaagc	96180
tttgggggtc	attatgcaat	gtatacatgt	gtgggtgcaa	gatgtaagct	gtgcatatgt	96240
gtggggggcca	aagggtctct	cctcaatccc	tctctgcctt	attttcatatt	aaattataat	96300
tattactatt	agtgtgtggt	gtgatgtgtg	tgggtgtggt	aagccctcac	ggcaatcaga	96360
ggatgtctgt	ggtctgagga	tgctctctta	ccatgttctg	gtgggttctg	tggatggaac	96420
tctgttagtc	aggtttgcaa	agctagtgtc	tttatctgcc	gagccacctt	gctggccttc	96480
aaccttattt	tttgcatgga	acatggaact	tcctgagttg	cctggacagc	aagtcccca	96540
gacctctctg	ttcctgcctc	cccctgtctn	nnntcacang	aggacacacn	gcttantggg	96600
tnctcggatt	gctgcncacc	tccccgcnc	ccnagcctcc	tgctcccccg	cccctcgccc	96660
ccgctggnc	ctcccccccc	cccccccccc	cccccccccc	cttccccccc	ccccnnnnnn	96720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	96780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	96840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	96900
agctctaact	gctgttaaat	tcacactcct	tcacatcccc	acgctaggac	tctaaggagg	96960
caccagcaag	gccaggtcc	agcttgactt	agagcaaagc	atcctcccc	ctccacacaa	97020
tggaaacgga	cggaaagggg	catggaagca	gaaccagaca	acagcagcct	agccaagccc	97080
aggactctgc	tccttcccc	catgcctgcc	gtgcaactgg	ggaggcaaa	ccccagccgg	97140
tgctttctga	cgcttagcg	gaagacaagg	ggagcctgtg	attatgattt	ctgctgattt	97200
gcaatgaaac	actaatgcag	tgggcttttc	attaagccag	atttattcaa	tctaaagatt	97260
ttatttcctt	tatgtagaaa	gtgcatcttt	atatgttgtt	ggaggagcag	agatgtgata	97320
aaaagaaatt	tctcttatga	actaatagca	ctgatacata	gtggtagcta	tgccataggcc	97380
tctctctctc	tctctctctc	tgtctcctgt	gcattgtgtg	gtgtgtgtgt	gtgtgtgtgt	97440
atgaatgcac	acaaagtagc	cccccccat	attatctctt	ctgtgggata	tccagactca	97500
gcaaattggtg	gtgactggga	agcttgccca	tgcaattctt	gccttttctc	ttgccagccc	97560
aatccctttg	cattcaaacc	cgggctgctt	gctgtggcca	gccctttcac	ctggagtcct	97620
tctctctctc	tacctgtctt	cccatccttt	gcagacaatt	atcctcaata	actagccaat	97680
tacccttaag	gacaattata	ctcttccatc	agcaaacacg	ggtgttcttt	ccttgagtct	97740
tttgatgaag	tcgataattaa	agagatgctt	tatttacata	aagtcaata	gctccctttt	97800
agaagggttt	gggttcgatg	tcaaaagttt	aaaatcttaa	ctagaggatg	ggtgtagagg	97860
gcttttggct	agggtagaaa	agagatggag	atacttatct	tgatgttgct	ttaaaaggta	97920
ggatgccag	agaaggtgga	aggatggggg	agggagggtc	cctctcaag	ctaataaatc	97980
taaaagcagg	gatgagctgg	gcgctaggag	tggaaaccagt	cagaagtgtc	tgcccttgagc	98040
tgaccacagc	tcttgccctc	ccctccccc	gtctctctgt	gaaccgccag	cattaggagc	98100
taattcgctt	agaagaccag	attggaatgt	gtgtctcacc	ctccactgct	cagaaaaacct	98160
ttattccagg	caaggactga	cccaaaccca	tcattggcatc	tgccaatcag	gaggccaaag	98220
gtgcccggcag	ggcgggacct	agctgtgcag	aaacagctcc	gttatggcgc	gcagaaaaag	98280
ctgggggggaa	aggctaccgt	tttatctctt	ggcagatggc	ttctctcttt	gatgcttttg	

gccttacctg	ttactgcctg	cacttgactt	gacctaggca	aaaatagcag	cgagatacag	98340
gttctcgaag	ttagaaggaa	aaaaaaaaag	ccccaacca	caacacaacc	cggaagtgtg	98400
ccccgcgtgt	gtttctaaag	agctgttttc	ttcccaagct	ctacagcgtg	gtggctctaa	98460
tcggaatttt	cttttttaac	atagcaggag	tcccaattag	cgtgttggtg	aatctttcaa	98520
gtagagtggg	agttccgtgg	ccacagagag	cagaggcaat	attcagcata	aagccctaga	98580
gaaagaggtg	ttgtgggcct	gtgcacacat	gtgtgtcaac	gcacatgtgg	cttgtggagg	98640
ctggcttccc	actctcaaga	tgaggtgtgt	gcacccagg	ccttttgatt	ctcaaagctt	98700
tattaggacc	agagggactg	tgtgtgtgga	ggggtgttgc	tcacagtgca	gaaacccaaa	98760
cctggcttct	ccaggagccc	acatgccaac	aaacaggctg	cacactcttg	ctagtacatc	98820
ccctaaaggt	atggggatga	gggaccaagt	gctttgcaag	acagcaggca	cagagttctg	98880
ggacgctcct	gtaccccaga	ctcagccgct	acccagggcc	agctctgac	tggtctgacc	98940
tactttcttc	tggtgtgtgt	tttggaagtg	ctgatgtcaa	tgagaatttc	agcagagtgg	99000
ctgagtgaga	aaaaagagga	gagggaggaa	aagggggggg	ggacgggacg	ggccgaggcc	99060
aacaggaaag	ggcaggcaac	aagacaatga	ccacaaggct	cctgtaacta	cactaactgc	99120
ttacctttcc	tgacccccag	ggcttagcca	atatagctga	gacccagtct	tggtgctgtg	99180
gcttcaggct	aagtaaacag	ggaagagtgt	gacatgggtc	tccattctct	ctcctcatcc	99240
aacaagggga	ggaggcagtg	gccaggcagc	gatgccatcc	gatgccatcc	ttctgggagg	99300
agccagacat	ttcaggcacc	tctccttccc	tggtgtccta	gaggtgctgt	gtctgcatcc	99360
atctgccatg	cctgccatct	gagagaggcc	actgggactt	ggtagagagg	ttctccacac	99420
atgctggcct	ggaggaaatt	ggtcttttag	gacactgaag	gcagtttcc	ctgttcagtg	99480
gctccttgga	aacccacgtg	acagagctcg	catgacaact	tgccggctct	caactcccat	99540
tcttagctgc	ctcaagcact	gtaaggttta	ggagagcccc	agatgtaagt	atggatggga	99600
agaccctcca	gggagtcatt	gacctaccct	ctgaactcta	acatgggtcca	gcttttccat	99660
tccacaattg	aggagacgcc	agacctggca	ggggagcaag	cctttgtttc	tgacccattt	99720
gcaaacccca	gccactgagg	aacttgcata	caagaaactg	cctctggggc	tctcctggac	99780
tgagccctgc	ctcccagggg	acaactgggc	aacagatcct	tccaggtggc	tgacgtgaca	99840
gatccatgct	tttatgacat	agaaaggcct	cagtctcagg	atttcacaca	ctgtattttc	99900
ctcatcctgg	ggaccaggga	aggcgagcat	cttctgctcc	cccaaaaca	gtgtgggaat	99960
gattaaaatc	attttttttt	tctgctccat	gaactccatac	agttttcaga	taccgaggag	100020
acaaagccct	cctgtgctga	aattagaccc	cgaaaaatag	gttagctgac	aattacttgt	100080
ttctaagtgg	agtgtgatgt	agtggcagga	gcgcaggatg	ggctgccagg	gctgcagctc	100140
cccccccccc	aaacttactg	tctcttaacc	tctcgagtcc	ctgggtttct	tgccgggatg	100200
ataattctcc	ccatctccct	cctctggtgg	gctggtggaa	agcgtaatga	atcaacgctt	100260
gaagcacgct	gaagaggcca	gactcggtat	gccatgtaag	tacacagcat	cgccagccac	100320
ctctcaagtc	tacacggagc	tgattttatt	acctcccgtg	aaagagacaa	caatcatcat	100380
atttacactt	catgccgcag	cttctgctgt	ggcacggcag	cacccctccc	ctctccgctg	100440
ctgaggactc	catcaagcac	gctgccttgc	caggatgaca	gcagcccact	ctcagcctct	100500
ccctggcctc	cttacagatc	atgacctcct	gccccgtgag	gtctgtcacc	cgaaaaccac	100560
ggtacaccgg	gggctgcagc	ctctctatgg	gggaggctga	ggaaatgaat	tccgtaggtg	100620
aaaggcttcc	taggaaatca	gacgttgcta	gtaattaaag	agcgaagcat	aggtgcgtga	100680
aaggtaaatg	gatgtttatt	aaatgttgcg	tcatttaaag	agtgtcctgg	tgcttcagtt	100740
ccttggttacc	atgcagggct	gtggacgggt	ggcaattagg	ctggcacggg	tagagctcac	100800
ctgctgagct	gagggagggt	ggggacacac	cttccggtaa	ttgctgctgg	gcagctctgg	100860
gtctccccac	ccccgcccc	gccctcactc	cccccccccc	acttctttcc	tgacagctct	100920
ttcatttgga	gcagcttaca	gggcttggtg	cccttaccba	gaaaatcacg	ttggaagaaa	100980
tataagaaaa	agaggaatga	aagagaaagc	cagaaaagtt	catattaggt	tcggatctgc	101040
ggccaaacct	ggccgagaga	atccatgacg	gtccgcgcgc	atataaccct	gtggcaacag	101100
ggcccgccac	aacaggggcc	gccacaagag	cttcttgagt	tgccacctgc	caggagacag	101160
gatgaatgaa	tgatcatctc	gtccttagag	cacaagccag	gcctgattct	ccaatattga	101220
tgtgtgaggg	agatgtcaac	agaggttccc	taaagaatga	tgcttctatt	tccatgctaa	101280
tectggggcg	tcagcttcag	tcggaacagc	cggaccgtta	ccttagctct	gctgttctcc	101340
tgtctgtaac	ccgcagaggg	aagggcgggg	tcacccagca	ttgccactcc	ccccaccctc	101400
acgtggtcca	gacccctctt	gggttgatct	gctcctgaaa	aacagtgttg	gctcaagttt	101460
gcctctgaag	gtatgtcacc	gctggctcag	ccagcttatc	tccccggtgc	tttcaagatc	101520
aaaacaccba	aacgaaagaa	aaactttggt	tcaagagcag	agtgtgggtg	caactctgat	101580
caaagtgttt	ttcagcatga	caactcactg	ccggtgacaa	ccagtacttg	gctgttgtgg	101640
ctcagagtga	gatgcggagg	gaagtggatg	acaacagctg	tatccagggtc	caaacagagt	101700
agattcacgg	ctggcagaaa	atggctgaga	gccttgggct	gcacccctcc	tccctcctg	101760
cctctctctc	ttttcaaggt	ggtttttggg	aatgtccttc	ctgtgggttg	tggtgccttt	101820
ccatgtagga	cctggggcct	gtgcagatgg	ccctgtgttc	ctgggtgctg	tggtgagatg	101880
tgaacgagtg	ataggaaccc	aggcactaaa	cacacaatgt	ggttgatctc	gactagaagc	101940
aaggcaagag	caggaggcat	ttgagggtaa	aggagtgtaa	ggactgtgta	aagagatgag	102000
ggttctatct	gggaggcagg	agtcccaatg	ccagcaaata	caatggactc	tcttggtcga	102060
cccaaccaga	gagaattcaa	gatggcagag	ggacaggctg	tctgagtttc	ctatggctgc	102120

accgataaat	ggctcataagc	agagtagagg	aaaaccacag	acagaaattc	atgccattga	102180
gactagaaat	ctagctcaag	gttgtgtgtg	gcaggggttg	ttcctgggtg	ttcaaccttt	102240
tcacactgtg	acatgatgct	gtgggtctgca	gatgtgttg	gctgcatcca	tagctaccct	102300
gggacacatt	catggaccgc	aggttacaca	tgctatttaa	aaactccaag	ggaagggcta	102360
gagaaatggc	ctggtagtta	agtatgcttg	ctgatcttcc	agaagacctg	agctctgttc	102420
ctagcagcca	tggtgggagc	cttacaacta	actatgactt	ctgagctcca	aagctctctt	102480
ctaatacata	catacatata	tacatacata	catacatata	tacatacata	cgtacacaca	102540
cacacacaca	cacacacaca	cacacacact	ttaaagaaaa	aaattctggg	ttggagaggt	102600
ggctcagcaa	ttaagagcac	tgactgctct	tacagagggtg	ctgagttcaa	ctctcaacca	102660
catggtggct	cacaaccatc	tgtaatggga	tctgatgccc	tcttctgggtg	tgcgctgaag	102720
acagatacaa	tgtactcata	tacattaaat	aaataataa	gaaagaaaga	aagaaagaaa	102780
gaaagaaaga	aagaaagtgt	aaacgaggaa	aattcctaata	taaaaagaa	agaaagaaag	102840
aaagaaagaa	agaaagaaag	aaggaaagga	attctgaggg	agaatctgcc	ccttttccta	102900
acttccaggg	ctataggcaa	cctgtggcct	ggggaagctg	tagacaacct	gtggcctggg	102960
gaagctgtag	acaacctgtg	gcttggggaa	gctgtagaca	acctgtggcc	tggggagagt	103020
gtagacaacc	tgtggcctgg	ggaagctgta	gacaacctgt	ggcctgtggc	agcatcatgt	103080
caacgcctca	cctctgtgac	ccaatttcc	gttctctaag	gacacatgcc	atcaaatagca	103140
taggacactc	tacatcaaga	tgatcttgct	tcaagatgtt	taacaaaatt	acatctgcaa	103200
agacctatct	ttacatgtga	ggtcactcca	caggttctag	acataatttt	gaggagccac	103260
catccaactc	actatgtgac	agagtcactc	agagatttgt	gtccaggaca	gactggctgt	103320
atctgctctg	agagtcacct	gcctgcccgt	gggaacctcc	cagtggctcc	taagggccct	103380
gaggactttg	gatctgcaaa	gccacatctt	ccaaaacctt	tttctctttt	tggagagcta	103440
ctctaccctg	aaaccttttt	ctctgagggtg	gcttttagag	aggcaggctc	cagcagggca	103500
ctgtgcccac	aagaagtccc	ggggagaagg	gacccaaggg	ccagtgtctga	actatcgctg	103560
agactgagaa	cattgtgtct	cacctaaaat	cgtgtgtcgc	aaggaccaag	caggctctat	103620
aaatgtctta	ctgcctttat	tccttttcc	ccgctccatc	ttactcctca	ttttgtttg	103680
tttgtgtgtt	tgtttgtttt	cttctgagat	tgagccagg	ctggccttca	gctcactatg	103740
taactaagga	tgactttaaa	cttctgatcc	ttcttccct	ccacttccag	agtctgtggg	103800
cagggtgtgtg	ccaccgtacc	ccagctttat	ttgagactat	gattcaggct	ccatacttca	103860
tgcatattag	gtaagcatgc	taccaacttg	gctatattcc	cagcctttct	ttcttctctc	103920
tttgagacaa	tgtctttttt	tttttaatta	tatgagtaca	ctgtatctgt	tttcagacac	103980
accagaagaa	ggcattggat	cctatttagag	atggttgtga	gccaccatgt	ggttgttggg	104040
atthgaactc	aggacctctg	gaagagcagt	cagtgccttt	aaccgctgag	ccatctcgcc	104100
agtccttgag	acaatgtctt	gctatatggc	acataattggc	ctcaaactca	gaatccttcc	104160
gcttcagcct	cctaaatact	gggattacat	gtgagccatg	gtgtttggct	tctagccttt	104220
cttccctccc	tttcccttcc	cttttccctt	ccctttccct	tttcccttcc	ctttccttcc	104280
cttcccttcc	cttcccttcc	cttcccttcc	cctcccttcc	cttcccttcc	cttcccttcc	104340
cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	104400
taactcctggc	tgtcttgga	cttgcctgtg	agaccaggct	ggcttgnnnn	nnnnnnnnnn	104460
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	104520
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnaaga	agaagaagaa	gaagaagaag	acaacaacga	104580
cgacaacacc	ggcgccgctg	cctccactgc	catccacctg	agacaggact	caaataccaga	104640
ccaattttta	aaaccagtgt	ttcaagccgg	tactactgaag	tagtagtccc	acttgggatt	104700
atagctcctt	actttgtttt	gctttgacgt	ttcttgacat	ggtgtgatgt	agctctggct	104760
gtcctagaac	tcaatgtgta	aattaggctg	gccttgaact	tgcttctgcc	tcctgtctggg	104820
atgatagact	gatggtgtaa	aactccactt	aggaggcaga	ggtgggagga	tcagaaattc	104880
aaagtcatcc	ttggctatgt	tgtgagtttg	aggaccaacc	ttggctacat	gatatacctat	104940
ctcaaaaaga	aataaatgta	ttgccgggca	tgggtgcaca	cgcttttaata	cccagcactt	105000
gggaggcaga	ggcaggcaga	ttttctgagt	tcgaagccag	cctggtctac	agagttagtt	105060
ccaggacagc	cagggctaca	cagagaaacc	ctgtctccaa	aaacaaaaaa	caaacaaaaa	105120
agtgaacccc	aacagtactg	ccggacagtc	tggtgtcttt	cctaagtctc	ctttcaactc	105180
tgtttaccca	ggtgtaccca	caaggtgtgt	gagcagctct	ataccagag	gtgatacggg	105240
tgtttgaaatg	agagaaaagt	ttcccatcag	ctcgggtgtg	tgaataactcg	gtccccagtt	105300
ggcagtattg	gctggagagg	tgatggggag	gtgtagcctt	ccggtggaga	tgggctttgg	105360
gagtttaag	cttccctccac	gaagaccact	gggtgagctt	tgttaccaac	agaattggat	105420
tggtcctgctc	ttggctccgg	cctcagttta	tctaaaattt	acatgttacc	tgatcaaaaa	105480
ctgtttcctc	ccccacccct	ctccctgtct	gtattccctg	ccctctagt	gtgctggctg	105540
tatactacac	tggtgatctt	gactgtat	cagtttaact	cttgttctct	ctctgctgac	105600
tctagagatg	tcctttcatg	gctgggacct	ggctcagaaa	tttctaaggc	actgagcctt	105660
cctccatctg	aacttttagga	aacttcttgg	cctaagggtg	tatttctgac	ttagtagtga	105720
atgagacctt	ggagcctgca	ctttgttgaag	ccacctgggt	ggtggtgggtg	gtggtggtgg	105780
tggtggtggc	agtagtagaa	ctctgatgaa	cagttagtta	ttcaaggccc	atctagaaaa	105840
agaaaggctt	tgggtgcaca	tggctataac	tcagtggcga	gagacgtgct	tcccatgaat	105900
aataatgatg	acgatgaaaa	taattctctg	tgactgttct	ccccacttcc	ctctctctca	105960

ccctagctct	tatctaccga	atccctgcac	aagcaccat	ggggtttaca	gaatctgggg	106020
cggaacgtta	gtcacttccc	ttcgccctact	tcagtattgt	gtttccagaa	gtaccattt	106080
tggctagtca	ctgaggaaaa	cggcagctgc	ctgtgggcca	ccagcccatg	ccaagtgagg	106140
tcagcaagaa	agaagctgac	agcaaatgtg	ccaactgttg	gtctgctgga	tttctactgt	106200
gctaagtgg	ttcaagaagt	ttcttcttaa	ccccctacaa	gaaaccacaa	attttattat	106260
ctacactgtt	ttgtagatga	agaaaaacacc	attccgaagc	tcactgccag	tcagcactgg	106320
aactggaatt	tgtttggcta	attcagtggt	tctcaaccag	ggtggatttg	gactccccag	106380
gggatatttg	gggacacttc	tggctgctcat	aattgggac	atgtgctact	ggcacctagg	106440
gtagaggcca	gggtggtact	gaccttccta	ctatgaacag	ggcagccatg	tacataaatt	106500
ctctcgatca	aaacatcaac	agtgtctaa	ttgagaaatc	tcaggctgaa	acctgtcat	106560
ttggccttga	ggtgggtggg	aggaggttag	agagtggaa	aaaatcagaa	gggccaccac	106620
agaggcctcg	agttagggag	gaacagggct	cctatgctag	ggataatgga	gaatagggca	106680
gcttggtgaa	acttttcttt	cttccaagct	tggctagagc	cctgctcaat	ttcccccaac	106740
tctgccaaagt	cagtccccgg	acttcgcact	aagtttgtcc	tggagtgacc	ctgactccag	106800
cttgaggctg	gggcaacaca	tttacctggg	tcttcccagg	agtgggttaa	aagtcaaaaga	106860
taagtggcat	ctgagaagtt	aaaagtgggg	tgaagtggga	aaaggcaggag	gaccccatca	106920
atcagctgac	ctaggaagga	agagagcaac	tgaagcagca	aagagctggt	ccaggagact	106980
ggattctgac	ccactaggct	tatcttccac	agcctttctt	gtttaggctt	gggctcagtt	107040
tccctgcatc	atccgcagga	gcccctggag	caccacacatt	cagccggccg	ccaggacagg	107100
ctccccagca	gtggcctccc	caactaactga	cagtgggtgac	aggaaatata	tccatttcca	107160
atctcctcag	aagtctgaat	aagtaaggga	cagatgttgg	ggagaggcgt	cactcttggg	107220
ttgatgaaga	aaagatcatg	agaagcatac	attttacccg	ctatgggttg	ggttccatta	107280
ccacccatgg	cggggttgag	ggggaagggc	agaaaaaagg	agatggagaa	ggacagacac	107340
gtaagaagga	ggatgtggtt	agggtctgatt	cgtcccctgg	ggtcgaaaca	tcagctgtta	107400
ctggggcgga	aggcaggaag	tctccttaaa	gacacaatat	tctgaacgtt	gaactcagga	107460
tttgaagcaa	gcccagcagt	caccttagtg	gagcccatat	ttaatttaac	acagaagcgg	107520
ctatctcagg	cttccctctc	atttctttgt	ctcagatgct	ctcacttaga	aagcttagat	107580
gctcttagaa	atgactcaaa	agtcaagaac	cccaggcca	aagtttctct	ttgggggttg	107640
ggagggagtg	aagaggtggt	cccagtcctg	tccctttaaa	taagcaattc	agcagctttt	107700
gccaaagtcat	tgggttcatt	tgggtttttg	cccattcccc	gcctttcaga	ctctgatttg	107760
cccctaggga	aggagccgcc	tcttcatttg	tctccacctt	tgaatcact	tccctaagta	107820
ggcctgagtc	agagaagcgt	ttcggagggc	gggactgaat	gggtgttaat	cttagaacccg	107880
ggtttctggt	tgatactact	ttggtaaaga	tcttccocta	atttttaaaa	agacgcttcc	107940
tctctaaaaag	tgagggcgaa	tcctttgtta	agaacgtgcc	ccttgagaag	ccgtgggctc	108000
ttcagcgact	aagacgagac	attcactaga	aaagatttca	ctaaacccac	gagggataga	108060
ctagacctcc	agtgaagatt	gggcctgtgc	gggtgacatt	tgtccctata	ccccgaagac	108120
ctcgagctag	ctctccagtg	aagactgggg	ccgtgcgagt	gacagtggtc	ctatacccc	108180
gaaaaaaaaa	aagtcctatt	tgtggaaaaa	aaaaaagact	tcgggtgttc	tgctgcatcg	108240
gtggcttggt	tccatcttta	gttctactca	ctcctgttgc	ttcgcgtgct	ccaccttcgc	108300
ttagctcagg	cctcctgtga	atcagttttg	aggctaaaag	aagttccaag	aaggaggggc	108360
tgtagccctt	taaggacttc	cccgcgaccc	agtcagagat	cagtttaaaa	atgccaaactc	108420
acagagcgcg	ctgcattctg	ggaagctgag	tgtcaccgta	agaacttcat	tgaccggaat	108480
gcaactgaaa	aatacacgcc	tatacttctt	tctgtctctt	aaactgtagt	ttgacgtaaa	108540
gctgggtctaa	gcaagtgcgc	taggcgaggg	gttagccaca	ccttttcagc	cattggccag	108600
ttgggttagtt	ggttaggcgtg	gcttagagaa	gctcctccag	gcaagggggg	ggcctccttg	108660
ccaatcagag	cccagacgcc	tgaatgggcg	ggagtaagca	gaggtgctgg	cgcccccgag	108720
tgggtgtggt	cacgttgccc	agcaatgggc	ggtgattggc	cctgggtggt	tcattcgcatg	108780
ctcgtgcgtc	acgacgccgc	cagctgatcg	gagactggag	ccggtgtgtg	ctgggcgctg	108840
ggaagagaca	gagcggctcg	ccgtgcggac	aggtcgcagt	gattttgctc	ctctgtccac	108900
agcaaccccc	gcacccagca	tcaggtgggt	gtgatctggg	gacccggtea	tccggggggg	108960
aaccgcggta	accgggtgat	ggggaagta	gggtcctgac	ggccacaccc	tgcccttctg	109020
ggggagggga	gagggggcgg	cggggacagg	ggcgctcttg	ggagaggagc	ctggactctc	109080
ccgagtagtg	tgtctggacg	tttaaaagaga	gagtcgccga	caggagtcgt	ggcagaaggt	109140
ttggagaagt	aactggggag	gaatatgaga	ggccagaggg	ccggggcggt	ctaaccgccga	109200
cgcccttttg	tttgaggatg	cccagctga	ccatttagcc	tagggaggat	ctggacgagc	109260
gaggggtgcg	gaggtgcatt	gcctctaccg	gcgctgactg	ggtcagggcc	agttcaagtc	109320
cctggcaggg	aaggggtcgc	tgggcggtcc	ggccctcctt	ctcgttccct	cccggggatg	109380
ttatgtaagg	ggggagggga	aaggagtagg	gggcggcggt	gcgaggccct	tatgcaaccc	109440
aaagggttagg	gtttcaccgc	gggtttggcg	gaggttgggg	ggggcgagca	ggaggagtgc	109500
ctggaaactc	taccgcaccc	cccctccca	gcctaacctg	ctgtcttgga	cagagagaag	109560
gtcacctttg	cacctcccc	ctagtatgtc	cggtagagag	gccccctagcc	cgggcttggc	109620
ctgactgcct	gggaagccgg	ctggctgggt	ggggcgctcg	ggttagtcat	cgctgggctc	109680
cctctctccc	cacctctcgg	ccaactcttg	gccccctccc	acggcctccg	gttaggctaa	109740
cgttcccacc	tccctctggc	cctagtttca	gtctccaact	catttggcct	gtcacctcgg	109800

ctgttagagt	aggctagaag	ctgtcatggt	gccagagagt	tgatggagca	gctggtcaga	109860
gggtcagtgc	cctgggcccc	ccccgccccg	cagccaaggg	cacctgcttg	gcacaaactc	109920
tcagcagcca	gtgaaccctg	tggcctgaac	agagctatcc	tgggcagaga	gaagtggaca	109980
gagactgatc	acctaggaga	aggaagatcc	gacaaagttt	atacttcccc	agaggctttt	110040
ggaatttgaa	agttggccac	cctagtgtaa	tctttccact	ctctgaaaat	agaaatcccc	110100
aggcaaaagt	tccttggccc	ttctatctgg	cagtggccat	gtccttggac	tgactgtgca	110160
gaaccaccct	ctcgggctcc	cagccctcta	gcctgccacg	ccccagccc	cctccctgag	110220
ccatgctgta	gggccccggc	ttttactgct	gattcatgcg	ttggaactgt	gggggcgggg	110280
cttggaactt	ggaacaaagt	tcagacgtgg	aggggcccgc	agacagcctg	gaattcatat	110340
cagatgtacc	cggaatgtgc	aagcggaaatg	cctggcatct	ctagtctga	ggaagctgcc	110400
cagccaccct	accataacct	ccctcccctc	cctgcttttg	tcagctgtcc	tcctcagac	110460
tcctgagagc	ccctgctgac	cttccaactc	tagtgccctt	cccatttcta	accctacaca	110520
aaccctcctt	gctgctgaat	tccctaagaa	caagtcatat	gagttgatca	cagagctcat	110580
atttctgaag	tacatttttt	tttttaactt	gggacttggt	ttctacaccc	tgccctttga	110640
atgccgaaga	tgtctgggctc	cttagcaggt	tgccaagagt	tgccagctcc	tagtctgtaa	110700
aggggacaaa	agcaagtgtca	tttagaagcc	tttgatctct	tattcaagaa	ccctcatta	110760
gaaggtactg	aaagtccagc	agagccaggt	ttggatggcc	tctgggtcgc	tgccctgtc	110820
accagctttt	cctgtttttt	tttttccctc	ccttcccttt	aggaacctgt	gcctcccaca	110880
ccctcacctg	gctgagccgc	agtagttcct	cagtggcaag	ctttatgtcc	tgacccagct	110940
aaagctgcc	gttgagaaga	tgttgccctc	tgccctggc	ttcgtggagg	aagaggagaa	111000
gcagcagctt	tgccatcat	ccggaaggtg	acagaactgg	ggtgggaagg	tctggacagc	111060
tggggtgatg	gctttatggg	agggaaacct	tggtcctctg	gggagccctt	acccccactg	111120
gcccagtgaa	agatttaggt	taaaggcact	gtctataaat	tggggaatag	gtgactccac	111180
ctccccaaga	ttagttagtg	tctgtgtggc	agtgggaaga	aatagaagga	aaagtctgtc	111240
tgtttactga	gacttccctg	taggcctgcc	ttcttatct	tcacatcac	catgccaaca	111300
cacacacaca	cacacacaca	cacacacaca	catacacaca	catacacaca	cacacacaca	111360
cacacacttt	cctttccatg	aggtccaaaa	gtaaatgtac	tcaggaaggg	ggacattgaa	111420
actccgttct	aagtagtcac	ttgtgtattt	actttttttg	tttatttgtt	tgattgactt	111480
tcgagacagg	gtttctctgt	atagccctgg	ctgtcctgga	actcactttg	tagactaggc	111540
tggcctcgaa	ctcagaaatc	tgctgcctc	tgctcccaa	gtgctgggat	taaaggcgtg	111600
tgccaccacc	gcccggctgt	atttacatct	ctttatttat	ttttagtctg	gcccagatct	111660
tgggttttag	ggtacttacc	cttacacctg	tgatttttcc	cacctgtata	atggggaatc	111720
ccatagataa	gtaggcagga	gggcattaaa	agtcaccag	tggtgactca	gagcctgggc	111780
tcttcttctt	ctcgtggatg	gaaacgaaac	agctcttcac	atgaactgtt	gtccttcccc	111840
caccccttga	ctactcacc	agctcagggg	gattaggatg	gaaggaaagg	ctatggttaa	111900
gtcccaggca	agctcgtggg	aggtagtcc	tctactggct	tctcaccatg	catgggtggt	111960
ccaaggcttt	ccctccacct	aaagcaaaac	tgtagctcct	ggttgggttc	tagcaaccac	112020
tgccatttat	tttctgcctt	tgctttccag	gatatgtaga	ctctgctcaa	tactgtgcag	112080
gcaagaaatt	gtcaggggag	atgggttgta	tgatatgagt	cccttctgct	gcctctagct	112140
cctgattcat	tctcacgtat	gggcttggtc	tctgattgtg	gttcaccttt	ggcccagctc	112200
tcctaacaga	agatgggttc	aggggttaca	ggaggctgtt	tgttgtatct	gacaggagga	112260
ggagttctag	cctgttcccc	atttgtgaga	aactgaaagt	cataggggag	actagatcat	112320
ctaataccagc	cccactgcag	tctaagctga	gggataggat	gtgtaaagga	ctgtagcaga	112380
cgggctgggg	aggtgagtc	ggctcacaca	ttgcgacaaa	gattgccctt	ccctgcacct	112440
cgcttgcttt	cttctcctct	cccttccctg	gccacagtgt	gtccctccag	cactgggtac	112500
atggctctgc	tgtcctcatc	caacatggag	cctcagaggt	gagaaagggc	agcctggaag	112560
caacagaggg	aggcacaaga	cagtggagga	cctggcctgg	aaccacaagg	gcctatccgg	112620
acattggtca	gagaggcacg	tagaagcctg	gagaacacca	ggaaagagag	cagccagcca	112680
gcctcagtga	aagacacgtg	cttccagcca	tctcctctca	ggacctgcct	tcctgggaga	112740
tgaagggcct	ccaggaagta	tggtcccatc	tctaccctgc	agtttctata	aacagcctca	112800
aggagcatga	gccacctctg	aaaggaaata	cacagcaaat	tcaaaaagag	attcaaatgt	112860
gtaacactgt	gggaaaacat	atctatgact	ggggttgtag	ctcagttggt	aggtttgctt	112920
aacatgcacc	aagccctggt	tctgtcttct	gcattgcata	aaactgaaca	ggttggccca	112980
ggtctgcaat	cccggcactc	tggaggtggt	ggcaaaggag	cctacattca	aggtaatcct	113040
ctgctataca	atgagttctg	agccagcctg	gctatataga	gactgtctca	aaaaataaaa	113100
caaaaataaaa	taaaagcattg	gttagtaatt	caaaagaaagc	agatgtggct	gaaaccgttt	113160
tcctgatca	taatacaaca	agcaaatgaa	agccagaaga	aggctcctgt	gcctgtgtgt	113220
tggcagtacc	aaccattgtg	agagatgcct	ttggacctgg	tagtttgctg	tcttagaaat	113280
gtatcctaaa	ataaggattt	ggttataaaa	tgttcatctc	agggttgtaa	tagagaaaaa	113340
tggaaagcag	ctgtttgttt	ggaagtccat	tcctttctgt	ctgtcatgaa	aatgtatagc	113400
tagggcttgc	ctaaagtaaat	tatatccatc	tgatgggtgt	gttctgtgca	gccatccaaa	113460
gtcttacaga	agaaaaattg	agtggaaata	taaataattga	atactaaaaa	gattataaaa	113520
gtatgagttt	gtgactgttt	ttaaaatatg	aacacatact	tgtaatatat	ttttttaaaa	113580
accatccaat	tgagtggaaa	tataaatact	gaataactaa	aagattatga	aaagtatgag	113640

tttgtgactg	tttttaaaat	atgaatgcat	acttgtaata	tatttttttaa	aaaaaactg	113700
aaagtggatt	caaaatgtta	agaatggttg	tttttgtatg	gtgggatagt	acaattgtga	113760
attttcccct	tggtttttct	gtctttctaa	tttttaaaata	ttgtgcattg	ctttcatatg	113820
ttaaataaaaa	tacaaaagac	aaataaatgt	tttaaaattt	ttactctttt	atgagtgttt	113880
tgccgtgtgag	tggcaggaat	ccaacattgt	cttctgaaag	cagctagcgc	taacttctga	113940
gccgtctctt	cactccctct	gtaattttta	aaaaatatat	ttgtatgtta	tattatgtgt	114000
ctttgtgcac	cagagtgtgg	gtgcacattc	tgcagaggcc	agaagagggc	atcagattcc	114060
ctggagctgc	acgctgtttg	gatcttctga	tgtggatgct	cagaatcgca	ctcaggctct	114120
ctagaagagc	agcaaagtct	cctagccact	aaagccatct	ctccctctag	tcctcattgt	114180
catgttttaga	ttttggagaa	tttgccttga	ggaggatggg	ctacaccaag	tgccagggtga	114240
aagaaaaatgt	ttgcttggga	tacctattgc	ttcttgagtg	tgtgtgtgca	tgcttgtgtg	114300
tgtgtgtgtga	tgtgtgtgtg	tacactggag	ctaggaatca	tatccagggg	ccttttcaag	114360
ctccaccaca	ctaaagtcaa	ttctatgaac	ttcattaatt	gtctgaatcc	acctactctc	114420
tacacacagg	aagcattcct	ctgactttct	gactgtcagc	cagctaagga	gggtgtggctt	114480
agaataagaa	agaagggaaa	tgctcaaaac	ctgtcactct	ntggggnnnn	nnnnnnnnnn	114540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	114600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	tattttacca	agcattgcag	gataaataag	114660
aacatagctt	aaagaattct	agagaaaaaa	aaaaaagtc	actaatgttc	aatgttttaa	114720
ttaatttttta	atgggaaaaa	gtaatatcaa	agaaatccat	gtctacggca	tattaatgtg	114780
gttaatgcaa	acaggacaga	ttatcttggc	caaattaaac	cacagtaaca	cgttttgaac	114840
taaaagtcat	ctattttatac	gagatgaaat	gtgaattagc	gtgcgcgcgc	aggtgcacac	114900
atgcacacat	acaaacagag	tctttttatt	cttaagccta	ccaaataaact	ctttaagcag	114960
taaataatttt	taagctctaa	aattttaaaaa	atagtgaaac	cccttcaggc	tatgacagaa	115020
tgtctgctttt	gccattcttt	ctgataaaag	tcccaaaggg	tgtcataatc	tgtatccttc	115080
ttagaaaagt	aaggagcaca	tcctatgagc	tggtagacctg	agttttacac	ccaagtcaca	115140
cgtcagcaca	cagcaatgtc	ctggaccatt	tgtgaggagc	cccacgctgg	tcctgagcaa	115200
caactcactt	ggactggtac	ggggtgaggc	tggcgatggg	caccactttg	gactgtgtcc	115260
cacctgaatt	ctgcagcagg	ccagtcccg	caggttttcc	aaatggcttt	gaggcaccat	115320
aagccttagc	tgcagtggaa	cctatgcaaa	cacacaggaa	aagggcagtc	attcccgttt	115380
tattttctcga	agtgaatgta	caggcaccat	ggcacgcac	tggaggtcag	atgataatct	115440
caaaaaaatc	caacttttgt	gagacaggct	cccttcattg	tcttcccacc	acacttgcca	115500
ggttaggggc	acaagctccc	aaggagtatc	tgtctccta	agggcactgc	catcagagac	115560
actctgacta	ctacacctgg	cttttaggtg	ggttctgagg	atttgaacgc	agatcatcat	115620
ggttgatatag	caaatttttta	ttcaccaagc	cagacaaccg	tttttttaact	tttttttaaaa	115680
aaagatttat	ttttatgagt	acagtgtggc	tgtcttcaga	cacaccagaa	gaagacatca	115740
gatcccatat	cagatggttg	tgagccacca	tgtggttgct	gggaattgaa	ctcaggacct	115800
ttggaagagc	agtcaagtgt	cttaaccgct	gagccatctc	tccagccctt	atttaacttt	115860
taaaagttta	aaattaattc	tcattatgca	cgtgccacag	ggaagtgtaa	cataagcaca	115920
tgtgtttgaa	gtatatatgg	ggatgtagtt	atgtgacggc	cagaggacaa	ctctgtgaag	115980
ttgattcttt	cctttcactg	tgggtcctcg	gagcgtcaat	agggtcagca	ggcattttac	116040
cagctgagct	atcttgtggg	ccaccaaact	tatttttaaa	aagctagaag	ttggttaaaag	116100
agaaggaatt	catgttaattg	aaatagaaaa	gcagagtga	agaggaggga	gaggagggtga	116160
gtatgatatg	gtaagaaaaa	ccaagaaaaa	acctctacaa	gtaaatctag	tcgtgttcca	116220
cagatcacta	tgttcaagga	tgggtgtccc	gaaaagggtc	ctaaaacaaa	gccagggaag	116280
taatcaaaat	gatcccatcc	agacctgcta	cagagaagct	ggggggaagg	agggagaggc	116340
cctgtaacct	caagtccccg	ggcagcacag	cagagctgtg	cccaatgatc	tccgtgatac	116400
aggacaacag	gatatctgag	aaggagtccc	agtgaagcgt	tagcagaagc	cagaggcctc	116460
gaacaaaacc	aacaacttca	ccatgatgaa	catcatgaag	atgtgtggac	aaaaggggtg	116520
actctgggac	acagtgtcac	acactacagc	ttccacctgg	gacaaaaagt	cacacactac	116580
agcttccacc	tgggacacaa	tgtcacacac	tacagcttcc	acagtcagat	tatttttctt	116640
ctagcggaga	ggctgcaagg	gtggagggca	ggtacaaggg	aacaggggtg	agtgggatcg	116700
gggtgcatga	tgtgaaactt	acaaagaacc	aataagttag	gaaagagaga	tgaagacata	116760
gaagccatgg	ccaacaaccc	cagcactctg	gaggcaggca	ggagggtctc	catgccaaag	116820
gtaatgaatg	ctatgtctaa	agagatgata	aagtcacatg	agagttacaa	agtcagagga	116880
gaaacttgga	cagaaaagcc	aaatgtaaca	aatgcatggt	aggtggggaa	aggtggggat	116940
ttaaaagaaa	tttcacaaaa	aggcacagaa	tgtaaaggta	aagcgagcga	ctaccagata	117000
aagcatcaac	taagaggagt	caggaacagt	gtatttttaac	tacctttata	aaaaaaatac	117060
tcgggctggt	gagatggctc	agcggttaag	agcaccgact	gctcttccaa	tggtcctgag	117120
ttcaaatccc	aacaaccaca	tgggtggtca	caaccatcca	taacaaaatc	tgatacggct	117180
ttctggagtg	tctgaaggca	gctacaagtg	tacttacata	taataagtaa	ataaataaat	117240
aagtattaaa	aaatatgcaa	taattggaat	tattctttgc	aattctacta	tgaatggagt	117300
ttttgttccct	gttgtttgtt	ttccgacct	gtttcatgta	atccaaactg	gcctggaacc	117360
tgctaaatag	tggaagatga	ccttgaactt	ctgattcccc	tgtctccacc	tcccaagtgc	117420
tggcctaggc	tactctgcct	gactcaggct	cagtgaattt	tcaaacactg	cttcaacctt	117480

gacccctaac	ttccatctct	tgggtcccat	cttaccatt	cccagactcc	cattctgtgg	117540
ctggggcttg	ctggctgggt	gtgtcgtgc	agaggctggg	gaagggactg	cttgcgtgctg	117600
ctgctgctgc	tgctgctgct	gctgctgccc	atatcctgga	gatcaaaaca	ggccaattca	117660
gtgcaaacc	agtcaagaga	tttcaacacc	agcagtaaat	accttaggaa	aaccacactt	117720
tggctcttag	aggagcagac	tgcaccacgc	tgcacaccc	gtgtctgcaa	agggtcact	117780
tttttgtctg	cgcatagtaa	agactgggtc	catttccctt	tcaactgttt	agaattgaca	117840
gctttcagtt	ttaatgaggt	tctccctgaa	gccttgccca	tttctccttc	aagaacctgc	117900
agtaaagtca	ctgggctcaa	tctgtgatac	taccactatt	tccatagcaa	caaatcgatg	117960
tcatcaacag	gacggtgtgt	tcgtggcagg	atccattaga	agagaaagca	gggtggtaag	118020
gaaaacctaa	aagcagttca	attgtctctc	aagtgccttg	gcttcaaagg	aaaaggagac	118080
gtttacaaag	ggttctaccc	tttcccccaa	acaaagcaac	cttattttgc	aactgacagc	118140
tgggtgaaaa	tatgagttga	aagagtctca	ggtaccgtga	agtgcctagt	aactgtcaca	118200
gcaggagag	atgagaacag	agttcagaag	cagccgtgtc	aaggcaacag	ggagataaaa	118260
aggaggggga	ccgagcacgc	ttctaaagca	gcaaataaga	cgcgccaccc	tgctggagtg	118320
tgtatttgcc	cttcaccctt	tatacaaatg	tttaccaggt	tgaagaatgt	taacattgta	118380
aattgcctac	cactattttc	aaaataacga	gtttatgggg	tttgatttta	cttctgtgat	118440
tggctctgaa	ggctgaagcc	agcctcggtt	aggctgctca	cctggagtgc	ttgtagatga	118500
caaaggcagc	taaaaaaaaa	aaaatgtccc	agagctcctg	aaacactaaa	actggtgcgc	118560
acaggcagga	agtctccctt	gccactgagg	ctctcccttc	tccaactgta	agtcttaact	118620
cctgctgggc	ctcctgaggt	cccaactcac	cgggcatgct	acatgcccc	ggaattcatt	118680
tttggaacaa	ttactcttaa	gtagctcaag	gaatgagacg	taattgtgtt	gctggagaca	118740
aatacattca	gcacaaagtt	gagaagatta	aatagaattg	attatgcttt	agttcatcct	118800
acagagagaa	aaagtgaagc	acctattttac	ataggagagg	ggcccaggct	actgacaatg	118860
agcccttctc	ttacagctca	gccttcccat	tactactcaa	ctcagacacc	cagccaggtt	118920
cagtctatgt	gaagcatttt	taaaatcagc	aagagaaagg	tgagttgctc	agtagtacac	118980
tgaaggaaag	tagaaatgag	cacacagtga	cctgctgcat	aagacacagt	ttaaaagggt	119040
acctaacttc	cagcgaagtt	cttgcccttc	ttaaaagaa	tgtggatttt	gctggtatgt	119100
gcacctcacg	aagggtgctg	cagctatgtg	gagggcagag	accatcctga	gggaatcggt	119160
tctccttcca	ccagagtcct	agggatggca	caaaggcacc	gggcttgaca	gatgccttta	119220
cctgctgagc	cctttcccca	gcctctcctt	ttctacagtc	tccaaattac	ctgaggtagg	119280
agcctcattc	tcagaacccat	ttctgctgcg	catgtctgac	tgaactggat	atggctcactc	119340
gactcaattt	ctataaaaag	ataactgagg	agccggcttg	gttggtagag	aaaacataaa	119400
gacccgagtt	ttgttcccac	atggggcggt	cgtgctagag	aggtgaggac	ggactaggca	119460
gggagcagag	gtaggcagat	cccaggcctc	aatagccagg	cagcccaacc	tgagctctaa	119520
gttcagtaag	ggagcctgta	accacagaag	acactgtgcc	gtgtaaccac	ggcaagcggt	119580
ccgctgcagt	ggcatctgct	catagccaca	gctacccagg	aggctgagcc	acaagactca	119640
ttgaagctgg	aggtcaaggc	cagcataggt	accataggta	gacccccatc	tcaaagttga	119700
acagtaaaac	tatatattac	tacaattaaa	aaaacaaggc	cgggtgtggt	ggcgcatacc	119760
tttaattcca	gcaactcgga	ggcagaggca	ggcggtattc	tgagttcgag	gccagcctgg	119820
tctacaaaag	gagttccagg	acagccaggg	ctacacagag	aaaccctgtc	tcgaaaaacc	119880
aaaaaaaaaa	aaaaaaaaaa	aaaaaaccaa	aaacccaaca	catgtaccta	ttctaacata	119940
aaattttcat	tttttataaa	aattacagtt	ataattttta	gttgaacata	atacaatgat	120000
aagtctccaa	cttgataatc	tgaggctggg	ggtgtagctc	atttagtaga	gtacctgcct	120060
agcagtcgct	aagtcgggtg	agtccttgcc	ctgtatccca	gtgcttgggg	aacaggcaga	120120
aagaggacca	gaagttcaag	gtgctcctcc	tcttcaggta	atgaggagtc	tgagccagcc	120180
tgggatgcgt	gagacacacc	agtaacagca	actactgcct	ggagtgtcga	ctctggacca	120240
ggaacaagag	ttaagtgcgt	tactcccgtg	actgtctgca	caatgatgga	gacagtacgt	120300
catctttaac	atctttgggt	gagaagagaa	aacctgggtat	tcctcagctc	gttctggcta	120360
agttcatgta	cttcatcatc	atgcagttaa	ttgttaaaac	accaatcctg	gcagcaataa	120420
ctctaattat	atagaacata	agatgtggta	attaggaaaa	gctactaatc	cacttaatag	120480
agtaaccttt	atctcttgca	aatctgggtac	aagagacagt	cccaaataca	atgatggcaa	120540
gaattccaga	gcattgtaaa	tagcagcaat	tgcccttcaa	ttaacgcatt	gtaacgcagc	120600
agctgcccac	aagacctcaa	atcaatcagt	ctatagctaa	ggaaaaatct	ttctaaagcc	120660
aaaaccattc	tacaaagcag	taacgtaggc	tccgtttata	ataacctggt	ttggggccac	120720
tggcaaatcaa	gctatcccag	gaagccagat	cgttaattctt	aggctctgct	ggctacacac	120780
tgggtcccaag	ccatgagggg	actagattac	agcaggctcc	gccctcggtg	acctgctcat	120840
agctatcatt	cttacagctc	attatggcaa	gtgagctctg	ggcagagaaa	aattcacaaa	120900
caaaccaccc	aacttcccaa	gcaagcattt	tcttaacaaa	cacaaagaat	aaataaatag	120960
agcctgccat	ggtagtgcac	acctgtaatc	tcagcatgtg	ggaggcagag	gcaggcagat	121020
ctctgcccag	agttccatgc	cagcctgggtc	tagacagttg	caagatcaac	aacgctatat	121080
ggtgaggccc	tgtctcaact	cccaaaccac	tgaaaacaag	taagacgata	tggatcaatg	121140
caatattttc	cagttcctac	tgacagaaaa	tggacacaa	taggctgggt	ataattccaa	121200
tatgaataat	aagtatatata	tacagtacct	agtttaagtc	tgagcaagat	attatcgcca	121260
acagcacaga	tacaggacac	acacacagct	gcaagcgtga	aggatttaaa	ggcaccatc	121320

cctacagtat	atgcagcatt	gactgtctag	ttttatttcg	cacactttga	atcatccatc	121380
cattttttgct	caatatggca	gcagtaataa	aatgtatatt	tgcattttga	tgcatgggtg	121440
aattcttaact	agcctgggct	gtgtttgctc	actaactcca	gtggacttct	gacgtaagag	121500
gcgctggact	agcttccaga	ggatgtaaat	ctaacctttg	ttctcggcct	cccctgaagc	121560
ctttgtctgtg	gtgaaagggtg	ctgtttctga	agccacgaca	gtcccatggt	ggtttgagta	121620
acaatactcc	tggtttggtg	acaatgccaa	aaaataccaa	aaaaacccaa	acccaaacca	121680
aaaaaagcac	agagctcaca	tctgagccaa	aaaaaccgac	actccctatt	ttttgaagaa	121740
ctcacagaaa	tcaagaagaa	aaacaagcaa	acaaacagca	cataacaggc	taacaacaac	121800
aacaagtcca	cttcagaggc	caaaaaacca	aaggcaaaag	ggctcttaaa	cattttacaag	121860
gacagacctc	actcagaaca	caagctatag	ccatgatact	gctcttcatc	catcaattct	121920
taaagacaca	gaagatgggc	tgctggcatg	actgggaaga	gaccagtgtg	ctcactcatt	121980
gctggggtac	gctgtaacaa	gcagaggaga	atgtctaaact	gtgtctgcca	atcacaggca	122040
cattcccagt	taaccagca	atcacttctg	gcaaaaaaag	cccactctgc	ccgcacactg	122100
gtgtacccag	gaggtaattc	actgcggtct	ctgggtggatc	ttttttcctt	cttgacagtgc	122160
ttgggttcaa	attaggtatca	agcacacttt	tatagtccaga	gcacggagac	aaaaggatct	122220
aaccacagag	ttggtttaat	aaacaacaga	acagccacac	cagaactgct	ggggggcggtg	122280
gggagggggg	ggagggaagc	gggagaagct	agaatcccg	cattcaggat	gcagagggtg	122340
gtaaacaggt	ttcaggctca	gtcgggggtat	aaggtaagac	tttatttcac	aaaaataaat	122400
aatttttttta	aagaggacag	aaaaaagaca	gcgtagagaa	ctagctatct	tttatgtaaa	122460
ccatgtgggg	tacaggagca	gcatttgcat	ttggttactt	ggcgaatagc	cctggtagga	122520
taaagaaacc	ttaaactgtg	ttatctataa	agggcagcag	tggtgtgcaca	atgcagtggg	122580
taggagccca	cccattgcac	gccatagttt	tgattacaaa	gtcacgcagc	tctactcaaa	122640
aattaaaaca	aacattaatg	cttggttaag	aaacagggtc	gcagagatga	tggtcagtg	122700
gttaagagca	cagcgcccat	atggaggctc	tcagccatct	gcttctccaa	ttccaggggga	122760
agctaaccgc	ctctttttggc	cttcaagagc	actgcatgca	catggtacgc	ttacctacat	122820
gcccgcgaac	attcaaagaa	aaatacaaac	tgctataaaa	cccataatag	accatcttaa	122880
gattctttca	ttttttttgag	acagggtttc	tctgtgtagc	cctggctgtt	ccagaacttg	122940
ctctgcagac	caggctggcc	tccaaccag	agatctgcct	gcctctgcct	ccacagtgtc	123000
ggaattaaag	gtattttaaca	cacacattat	acatatcttc	cttcttcttc	ttcttttttt	123060
aaagcggtttt	gttggtttttt	gttaagtttc	aattaaaaaa	ctacatagtt	ttataggcaa	123120
acataattaa	aaatgccaat	gtgaaataaa	taatataatac	atatataaca	ttctgtaata	123180
gattcactca	cacaacttat	atacttaaat	acaattttca	caataatgaa	aagctttgaa	123240
atgaagactt	ctggatacat	tagaaacgta	ccctgaaaaat	cgcaaatgac	ggttttcatt	123300
tctttgtgtc	agacattagt	gtgagtgtct	aaacttgcat	aaaggctctc	ttctctatca	123360
cttctactct	attgcagtgg	ttctcgctaa	acctggaacc	tggtgtcctc	gttctgtggc	123420
tggtactaaa	ggcagcaagt	cccagcaatc	ctcctgtctc	acctttcttg	gaaccaatgt	123480
tataagtgtg	tgtgggcact	agccttggtt	catggctgct	gggactagaa	ctctgggtct	123540
caatattagg	catcaagagc	tcttaactgc	taagccatct	ttctaccctg	attagaattt	123600
cttgaagcaa	aagaaactca	cagatgggtc	gagtttacac	acacacacac	acacacacac	123660
acatacacac	acacacacac	tcacacacca	aggcttagtg	accactgtga	aaagggaagt	123720
gcggtgagga	actgtaaaaa	taagggtgtc	ggaagctctc	gcacaaaatg	gtgtcctctg	123780
gacaagccag	ggcctctgcc	ctcatcagct	cttagtaact	atgggtgcct	acagcaaaac	123840
atgccaggga	ccactctaac	atggagctgg	gatgggctcg	agaggccctg	ttattaaagg	123900
agaagctgta	gagagttgat	ttgatggatt	ctaaagttag	gggaatcgat	ttcctttatc	123960
gatattgggtc	agccatgctc	tagttagtgg	ccccacaccc	acccatgagt	atgtgtggac	124020
agcacacacc	ggacctggca	agtcaataaa	acaaaaacaa	aaacaaacaa	ataaaagtct	124080
tggtccacca	tagtggctgc	tcgtgggttg	gagctgtccc	gcacttatgg	cagggaagaca	124140
gtggctgcaa	agaggacaaa	agtctctgga	actgatcaac	tctagagtct	gcttggtatg	124200
agaactggga	agtacccgct	gggacagaag	cagactctga	aggtgatcag	gacagagatc	124260
acagaaggag	agactgggtt	tcggaggaaa	tctgaaacat	aactcgacgc	atactgggtc	124320
aaactgggtc	ccatcactac	aacagcagta	attgaattgg	gcacaacatt	cagaaaaacag	124380
aaaaagacta	cagagtacgc	accctggcta	tcatacaacc	aggcgattct	ggcactattg	124440
gaagcaagcc	agactggaga	aaaggaaaca	aaaagttatt	taacaaaact	tcccagagca	124500
tggtaaaaaa	aaaaaaaaga	aagccagaca	tggtgggtgca	cgtttttaac	cctagcactc	124560
aggaggcaga	ggcagggtgag	gcaggaggat	ccatgagttc	gaggccagcc	tggtctgtac	124620
agtgggttcc	aggaaagcca	ggcaacaaag	aaaccctgtc	tcaaaaatca	ctgactgggg	124680
agagagggaag	tggatccaag	agcaagagag	agagagcaga	gagagtgggg	gtttggatgt	124740
cagcggttatt	aatgacagc	agaaaagatg	ggcccgacca	atgacatccc	agaaatggca	124800
aagatgaaaa	aataaacaca	agttctaaat	atcattttta	taatggctgt	gtgtctgctg	124860
gctcattctt	gttatccaaa	caactaaag	caggggtggc	cggtatttac	ggccagccag	124920
gactagagtg	agacactgct	ttaaaaaagc	aataagtgca	cgctaaccat	taatcagcgt	124980
aactggagtt	tgaggggagg	agtgggcccc	ggaagcctgt	gcctataaac	ccaacctgcg	125040
aggcctgaag	cccgaagggt	gaactgagaa	catctcaaga	caaagcacag	gcacaatctc	125100
ttacaaacag	tttaaacaca	ataccagcaa	ttaattgtca	gctttatgac	agataggctg	125160

acaggcatac	cacaaagatc	gggaagaaga	aacgggattc	ttgcacaaca	ttttacaaat	125220
cgacacagct	gagcctagtg	acaggccgtg	ataagctcaa	aacatgcaca	ctgcaaacac	125280
cacagcatgg	ccagagtgac	agggtcacag	caatacagca	acagagacag	taagaaacat	125340
tggaaaaggg	aagaaggcaa	agagcaccac	gagagaaaag	ttaacccgcg	attcccatatg	125400
agggcccagc	atccctctcc	cacctgacag	agaaaaccag	gaggacgcag	ctgaaacact	125460
gtacaactat	aaatactcct	tcctagtgta	gacagagttt	accaaaggg	atcgtaatct	125520
gaagcacaaa	cataacactg	taagttacca	gaagtttaga	cacgttgaga	attttaataga	125580
agttgagaca	tcaaacacac	gtttgatcct	aggggaatta	aattattggg	aacagaaaag	125640
ctctttacaa	agtctccaca	tttataaacc	aaatacactt	gtatcagaa	tgttgtctga	125700
gccagaggcg	gcggcacaa	ggtggctgag	gcagcaggat	caagtctgga	gaagcccctg	125760
ggtccaggga	gcactagtga	attcaagtaa	acattaaaca	ttaaataaga	aagacggatt	125820
ctacaggagc	ctagacaccc	tgtaaagtca	gtactactaa	gataccaaaa	ccaaactgac	125880
tcagggggct	gatgagatgg	ctcagtggtc	aagggcagtc	actgctcttg	cagaggacct	125940
gagttcagtt	cccagaaccc	gtatcaggca	gttcgcccac	ctgtaatgta	ccagagaccc	126000
taaacattttc	tatcctccaa	gcacagacac	acataaacat	aattaaaaat	aaatcttaaa	126060
aaaaaatctc	tcgtggacag	acagcaaaaa	atactgaacg	aaatctaate	aggtagactc	126120
agcaaagcaa	atccaaatat	gcgaaaagga	taatacaatg	caaateccaga	cttatcccag	126180
aaaccccagg	tcgcttttagc	atccaaaaaa	ttcaatcatc	ataattcacc	gtattagcta	126240
accgaaacag	aaaaggcatt	tgattattaa	taaatacagg	gaaagcattt	gacaacattg	126300
accatttcact	cttaataaaa	atgttaccaa	acaggagaaa	gaagtaaaat	ctcctcaacc	126360
cgatgaacag	gaacacaaa	tggttgggca	cagtggcaca	tcctgaaacc	ccagcaccca	126420
gatagctgcc	agactgggct	ggcatgagaa	cgaagggtaa	gaggcaccca	aacacttaat	126480
gatccaggagc	gtgcctccc	ccccccccct	gtgtcactta	aatccgtaat	taaaattctg	126540
ggaggagggg	tttgacacag	ggtctcactc	tgtggcacag	gctagcctga	atggagtcaa	126600
tgctggcctc	aaaactcttc	tgctcagtt	ccagagtggg	gggaaacaat	catgagccac	126660
ctcaccaggt	tttaatgcc	tttaacagta	caaggttaaa	cgctgcctga	aaagaacaag	126720
acaaagaaa	atcacacaga	caaacatcac	acaggtgctg	tttgagacta	ggtctctagt	126780
gatgtaggct	ggtctcaaac	ttgcacaaga	aaagaatact	tttgccatca	agatcaaggt	126840
tgctatcacc	catggtggca	tttcggaagc	agaggcagga	agatcagtag	tacaaggtca	126900
tcctcagcta	ccatgggtgt	gaggccagcc	aaggcagcac	gtgagccagc	tgtggaagca	126960
cacatctgta	acccaccac	tcaggagggt	gaagcaggaa	gttcaagcca	cacaagagcc	127020
acttaaaaa	aaataaaaa	ataaataagg	ttggagagag	ggnnnnnnnn	nnnnnnnnnn	127080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	127140
nnnnnnnnnn	nnnnnnnnnn	nnngaaaaca	cccattagac	tgctgaagag	actgagaaat	127200
tgctatcttg	aagggttgaa	atcactgagg	aatggcta	gacaaaactg	cttaaggaga	127260
attttattaa	aaaacaaaac	tccttctgga	ggaagtactg	gctcccttca	gcattaaacta	127320
ctgtgccagg	gaacccactg	ccagcagaca	cccaacttca	aggataatca	agcttgctac	127380
gtaacatggc	tgcttacgca	gacaccacac	acacttccctg	tatgctgggc	atcacctcta	127440
gcttacttat	aatatacaac	ataagtacca	tgtgagcgt	acagggcctg	gccaatgaac	127500
agcgatggga	aaaataacct	gcacatggtc	agtaatgggc	aatccactcc	ccagacactt	127560
ctaacctcag	ctgcagactt	gtgggtactg	agaacggacg	tgactaagtc	aattcacaga	127620
agcagagccc	tgtgcctgtg	ccagggaacg	gaggacgctg	ctcaatgggc	agagcttcag	127680
ggcgggaaga	gatggaaagt	ggagctaggg	cccagtgacg	tgaatatggc	taatgccata	127740
ttgtgtgtgt	gtgtgtgtgt	gtgtgtgtat	gtgtgtgtat	gtggtgtgta	tgtatgtgtg	127800
tgcatgtgtg	tgtgtgtgtg	tgatgtgtat	tgatgtgtat	caaaatgggt	ataaaaaatcc	127860
catacaatgg	ctatgcttat	gtgtctttta	ccacaagtaa	aaaattttta	gtaatcttag	127920
gaacacattt	ctaaaatttg	gaatacttgt	tggggaagct	atctgagccc	cagcatgaat	127980
caggaaatgg	gtaggagagg	caggagagat	ggcttagtgg	ttaacacaca	catgggtggc	128040
ctttcctggc	accaactagg	tcactcacaa	ccaggatctg	acagcctctc	ctggcctcct	128100
caggcaccag	gcaggaaaagc	ggcccccacat	ccctgcagggt	aaaaacattg	tactgaaact	128160
aaataaaaa	ggagcagaag	ctgggcctgg	aagtccctta	cctccagcac	cgggaagcag	128220
gtggatattt	tgtgagttcc	gggtctacat	agtaaaaaact	tgtcgccaag	taaaacaaaa	128280
caaaaactgg	gggctgggtga	gatggctcag	tgggtaagag	cacccgactg	ctcttccaaa	128340
ggtccaaagt	tcaaatecca	gcaaccacat	ggtggctcac	aaccatccgc	aacaagatcc	128400
tccttctggag	tgtctgaaga	cagcaacagt	gaactgtacat	atattaataa	ataaatcttt	128460
aaaaaaaata	ccttttaaaaa	aacaaaaaaa	aaactcccac	taaaataatt	ggaaaagtcc	128520
agggaaagcc	acactcgatg	gcgcaagcgt	gtgatcacag	tattctgtag	gtaaggcaag	128580
aggatcacag	cgggacggag	gccagcctca	gctacacagg	ccggtctggg	ctacagtgtg	128640
agaccccggg	ctcaaaaaca	aacaaaaaaa	aaaagcggtt	cacattatca	tatactcaag	128700
gccacaaaac	accttctctt	tcccaatcct	tgaatgtctat	cgaatttggt	tacttttttt	128760
tggttaattt	ttgtttgctt	ttcaaggcag	ggtttctctg	tgtagccctg	gctatcctgg	128820
aactcactct	gtagaccagg	ctggcctcta	actcacagag	atgtaattct	ttttcaaaag	128880
tggattttga	tttgggggtg	tgtgggtgat	accacaggga	cacttggggc	ccaaacccaa	128940
ccaaagaaaa	aacaccccc	ccccaaagta	tgaacatcaa	ctgtatataa	aactaacagt	129000

tcatataagc	taagtagcct	gcaggtacat	ggttatggaa	acagctttat	catacagact	129060
tcttttcagta	gatgccattt	gagaaaaaaa	aaaaaacaaa	acattctttg	gattttcaac	129120
atgtaaacga	aaatatgtta	aatcttaaaa	aaccttaaa	cagacgccac	tgctctttgg	129180
ggtcagggaa	gcacggctgc	aggcccccag	agcagcacat	tccctgaggg	aagccttggt	129240
cgtcctcacc	agcaccaaga	acagccaact	ataataagct	cataaaaaat	cctgaaaggc	129300
tgcccaagcc	tgaaaatctg	atatgaaaac	agaggccaag	acagataaaa	ggcaaacctat	129360
actttaaaag	tcaattccag	ctatctgtgt	ggaaggactt	ctcgggacgg	actatctgct	129420
aagaccttgt	gggtgattta	aacagcgtgg	agggcagaaa	aagaaaggaa	agtgcaggaa	129480
acaatggcaa	aagctcctcc	tccgtggcct	ctacacgcat	ccaagcgata	tgaggggagg	129540
agggaaaggac	actatacaca	ggtaccaatc	cccacgaact	agaaaaacaca	gcttttacta	129600
actagtctat	tttttttttt	accttagcta	gtgtctttct	tatgtttggc	ataatttctg	129660
ctctgcatta	aataaaccta	gtgtataaga	aggcaaatga	gtaaaaggtaa	agtcagttaa	129720
tgttcatttg	ttttgactgt	gtgggtgca	tgccaggcgc	attcgtggag	gtcagaggac	129780
aaccgcatgg	agcaggctct	ctccttcac	cacccagggg	gtcgcagggg	tcccagggat	129840
ggaactcggg	tggtcagatt	tgcttgcaa	gggcttcaact	caactgaggcc	ccccagggt	129900
cccaattcac	ctgtttactc	taatgtatat	atttttgaga	cagggtcttg	ctatgtagtc	129960
caggctagct	ttgaactcac	tggtgacact	agctggcctg	gaggtctgag	tgaacctcct	130020
gtttccaccc	aagtgtcggg	acaacagaca	gcaaccatca	aggcaaatg	aagtcttccc	130080
ttcacaccaa	agtggcttcc	tatgaccttc	tcctgacaac	cccaaacagc	aagtgcccg	130140
atatgcactt	gcgtttctct	ttttttactt	caagctttga	ctccactttc	ctaaaaagggt	130200
tttactagcg	tgaagcacac	ttcaggagac	acctgtagc	tctggagtga	ccggaacac	130260
accagctctg	atgcaaaaac	aaaaaacagc	gtgggggatg	gggcaagggt	ggcttgtgct	130320
tctgtgcaat	ggcattccct	gacgcattac	ctcacaccag	atacatacag	aaaagacaaa	130380
ggtaagtcct	cactcagtgc	gcgcattccac	cacacccaaa	gcaagcattc	aacagctcag	130440
tggcataagg	ccagcattat	cacagcgagg	acaaacacaa	agccaacctt	ctctttgagt	130500
tccccaagggt	acatagtgtt	ctgtccaagg	atggcctctg	cagactcctg	gaaacacgtc	130560
accactgggt	tctcttgaaa	atctgcaata	tttgccataa	aaacacaact	ctattatttc	130620
agcgaataaa	taacaaaagg	aagacacccc	aaggacgagg	gaagattatc	ttttctact	130680
taaaagtaact	tccaagggag	aaatgacatt	gccgccacac	acagcccccg	actctttgggt	130740
taacggttcc	cccttatgag	gaagtattca	cttctgggtc	ccccgtgttt	cacttttagt	130800
acagtatcta	atgaattaca	ggagccaactc	aataccttta	tcataaagt	ggctttgtgt	130860
taggtgattc	tgaccaagcg	taggctatgc	aaatgctctg	agcacactta	ggggaggctg	130920
ggcgtgcta	gccctgctac	atgcaggtaa	tgctacgcca	gtcttacgcc	acaggaagaa	130980
gcttaaaaatg	tggtcgaaact	caatctgcta	gagggtgtgag	tttagacctc	gggagacggg	131040
ccaagtaaac	ttccgcagat	gtcggagcat	cagacagagc	tgtcccctcc	tgatgcaaaa	131100
ggcttcgcac	ggcaagtgtt	taatgagcct	ccatggtaat	agtgggttcc	tctctctcct	131160
cctcccttat	caatatgctt	gacatctttt	agttttttga	gacagcaaat	aacgtagccc	131220
aatctgcccc	catacttact	gggcagccac	agctggcctc	gaactcctcc	tgctctcct	131280
ttaccccgta	acaagtgtc	gggttacaga	catgttcggc	catgttcagt	gtaagtgcac	131340
gctgcccagc	agggcaaaagt	cttccttatt	tacaaagcag	cagccaagcc	ctgcagccca	131400
ggcctatctg	atttctcag	cacaccccca	agggctctac	cgatggcagt	cagtccatga	131460
acaccgtagg	cttctcacaa	tccacactac	tcaccgataa	gatcatgcgg	tatttgaaat	131520
tgggaaatc	ccggtcacac	ttctcacagc	ggtacaaccc	attctgctgg	tcaatcactt	131580
tctattgtca	gtcctgggtt	gggcaggcct	ggtacataca	gttctctttg	cggagaagaa	131640
ccaccgctgc	cacagtgtc	aaatagtcct	cctgcaggcg	aaaggaagac	cgccatcagc	131700
aagcaacaca	ggtctggaac	aggcaactca	aagctgctct	tccttgcaag	tggtgagcgc	131760
gtgtacatcc	tagctccac	gctcacgagc	gatatgcaga	atcactaaact	ctggtttcag	131820
aaatcacacg	tgctacacgc	agaacccaaa	gaagtaacaa	accggcactc	acgcacctca	131880
ctttttccat	tttgagacg	gcggctcact	agctagcctt	gaactcagag	tttggctcgc	131940
ctctgtctct	atagagtgc	tggtaccac	acctggggc	tttgttttcg	agacaggggt	132000
tctccggaac	tactctgta	gaccaggtg	gctcaaaact	cagaaatcca	cctgcctctg	132060
cctcccaagc	gctaggatta	aaggcatgtg	ccatcagcgc	ctgacccaat	tctttttatt	132120
tatagttatt	atttgttaca	tatgtgtatc	agtgtgtatg	acgtccatat	gtatatgact	132180
gtgtgcagt	tgcatatgtg	tgcaagtcag	aggacaactc	tcaggagtca	gttctctcct	132240
ctactgtgg	cgtttgggga	actcagggtc	caagatagca	ggaaaagtgc	ctttacacagc	132300
tgagttatct	cgacactgac	atgtaatgat	tcagtgtgca	cacagtgggt	ttgcatgtag	132360
tcatataaac	cacagcgtgc	atgtgcaggc	taggaaacaa	cctgtgggag	ttggtgatct	132420
ccctcaccca	tgtgagttag	ggagaggaac	tcaggctgtc	aggctcgggt	gcagtgcctt	132480
tactcactga	gttcccttgc	tgcccaagca	ttatttataa	gatgggtgatg	tctaccctta	132540
tcttttaggga	tcagtaagtt	tttcccaag	acaaagccaga	aaaacctttt	aggcccaatg	132600
agccattcag	tcagtctcta	ctgcactcc	tcaacctgtc	tgtggggcag	gacaagccac	132660
agacaacctg	aagacggaag	gtgagccaat	aaaatcttac	tgacagtaac	agccagccag	132720
ctcacaggct	tgacaggcaa	ttcttggaact	gaatgcgttt	aagagaatgc	agaattaccc	132780
atgactaaga	tcttctaaat	ggaaaaatgt	ctggttaagt	aaccagcaa	ggagctaagt	132840

cacgcaagcg	gtggatacct	gctttgcctc	tgaaccctgc	acagggtttg	gtttatgttc	132900
aatcatgtca	agtacctaca	aatccagat	ctagcctgaa	ttcaaagata	ggctacctac	132960
cctgcccccg	gacccccacc	cggggtctca	ctgtgtagtc	ctggctgtcg	tagcgctctc	133020
tatacagacc	agctgtttatc	aaattcagag	accactttgc	ctoccaaagt	ttgggattaa	133080
aggcatttgc	cactatgcct	ggctctcact	ggttgtacca	gaggagcaaa	acaatgtggc	133140
cattttaaaga	gacgaagcta	gaaaccagtt	tagacagctt	tggggctgtg	ggtgtagctc	133200
agtggaaagag	cttgcttagc	atgcacaagc	tgttggtttt	aacctcagc	gtgacagaac	133260
cgaatacata	agaaggctta	gaggaggggg	tgggtgtgga	gagatgggtt	agaggttaag	133320
agctgggtgc	ttaattttccc	agtccccaca	tgggtggctca	caacatcca	aactgcagtt	133380
ccaggggatc	tgatgtcctc	ttctgacctc	cttggggaca	tgcgactcat	ttggctcaca	133440
tgcaggatgc	ctcggggccac	tatgctggct	caggggtgaa	ggtgcttgct	gccaagctgg	133500
gtggatttga	gtttgggtccc	tgggacccac	aaggaagagt	ttgacttcta	tacactgagg	133560
tggaaatgc	atgctctacc	cgcaaattaa	aaacttaaaa	tttaaagagg	aagctgtaga	133620
gaaatagctt	ttaggaggat	gcctaaggaa	cttctctgcg	ttttcagggtg	agattcagac	133680
tcaaagccca	attttaaagt	ttgagtgtctg	tcacgtgttc	tgtatgcccc	gttctggctg	133740
tccgttgtct	gtctttagtt	taagagagca	actgggtgag	aagtaactga	gagtctagcc	133800
gatgtttaat	tctcaagatg	tccgtgtgata	gcattataag	ttgctgtgga	tgacagtgtc	133860
ggatcgagca	cacagtga	tgagacagt	aggaagaaaa	cctatatgt	atacaggagc	133920
aataatggag	cagtggcgag	cgatattttg	ggaatacaat	agaaataaat	gattcaaata	133980
tccaagagaa	cgcagggttag	accaaggtag	gaagaggatc	actgggctgg	agagatggct	134040
cagtgtttaa	gagcactgac	tgtcttctctg	aaggctcatga	gttcaaattcc	cagcaaccac	134100
gtgggtggctc	acaaccatcc	gtaataatat	ctgatgccct	cttctggagt	gtctgaagac	134160
ggctacagt	tacttacata	taataaataa	ataaatcttt	aattaaaaaa	aaaaaaaaaa	134220
aggaagagga	ccactgagca	cacattgaaa	tggaaagtgc	aactgaaatg	caataaccaga	134280
gccagtctgt	ggcatggcag	caggaattac	agtctgttca	aaaccagca	cggacctgag	134340
gctacattat	gtccctctac	actggccgtg	actgaaaagc	agcaacgtgg	taccctggag	134400
caggcctgag	gtcccaggta	aggacacagc	cctgacctgg	cactaggaaa	caggtgtaaa	134460
aactcaagcc	cctgccgttc	tctgggtgtc	tggagcgagg	ggtgcgagg	acctgtctc	134520
cctggccag	gttctcagat	ttagcctcat	gcaaagt	ccagtgtgtg	ttgccccctc	134580
cgccccctcc	actcctgtgg	tcagagatgg	aaacaccatc	taaggcttgt	ccttctgagt	134640
caaacctagg	ggagaaagaa	acaaacgtac	tgtgtacagg	aacttcacat	ccttctcaaa	134700
tgtgtctgta	atgcaccagc	cccagagctgc	gccccctcgg	ctctcaccta	ggtctcagca	134760
cacctagtcc	actcaagaaa	tgcaatgccg	ttgtccttc	catctcgtgt	ccttcaattc	134820
ccctcctcct	ccaccggcaa	ggtcagacaa	actgccaaa	taccacagac	cctttcccta	134880
aactgggctc	acattgacag	cagccatgaa	caaaggcagc	aacagcaagc	cacactggca	134940
cccctcccc	caccctcggc	ctcaagcctt	caccccaatc	acaaaaatga	accagatgaa	135000
aagggacatt	tttgtttcat	gagtctgcca	aaaatgtttt	catgggagaa	aatcttttaa	135060
gcccaatcag	ttatagaact	caatcagaag	gcattatctg	ctgtgttaga	gatttgcatt	135120
ctcattgaga	atttgtttat	ggaaatatca	aggttaaat	catttatatg	aaattattat	135180
taaaattgac	tatacttctt	gacctctat	tggctaattc	gaaggaaaag	aggccaggga	135240
gaaggtagca	aggacaggct	agtggcagag	gacagtgagc	ctgagatgaa	aacacctcaa	135300
ctacacgatt	ggtggttaact	aaagctgccc	ccacacagga	ctgactccct	aaggagactc	135360
ccaacagagt	agcgggtggcc	caaggcaagc	agtcacatcc	gtcagaggac	aagggtccata	135420
aatccgtctc	accactacta	gccgtcgaag	tttaacttca	cacagcctc	ggcaaccaa	135480
gacggctcat	caggagtga	cagagccagg	agacagcagc	cctctcttat	agtaaaagt	135540
ccactcttag	agccaagatc	taaagatcgc	taatccttgc	ctgggtaggg	ggggtatact	135600
gaactgttag	atcaatgaca	gtattttgtt	cgggggacag	acatgcctgt	ttgtgatttc	135660
tgtgatcaca	gttactgaca	gtgaaatatg	aacaccttaa	tgtagacaca	caattaacct	135720
ggggttgtgc	tcagagtctg	ggagaaaaac	caaccaacaa	accagaccaa	tgccatggct	135780
gtctctgggtc	agacttctgt	cctacaatgc	aacgctagct	atcagcacag	ttacacagcc	135840
agaccacac	cactgaggat	gccttaagt	acagacaccc	cagatggatg	ctctaagtag	135900
taaatattgt	tattgtttaa	catcagactc	aaaacaaaaa	tgaggagcgt	gaaatatgag	135960
ctgtttttct	tgagtctctt	aaattctacg	acacagctgg	aaaccacaca	tgcccacctc	136020
ggtactacag	ctattcaatt	tccttaagct	tgggctaggt	aaagtatttt	ctttgaccac	136080
ctcaatcctt	ctacacaaca	cctcaggat	gaagtgtctc	tccaaggcag	agtttaatgc	136140
cttaaaaaagc	tagatatggt	gtggaggttc	taacaagctc	tcctcatctc	ctgagtgtct	136200
tgggaatgac	aatgtacagg	caacaaagag	tccattgttc	tcactgtact	ttaccgagtc	136260
cagtgaacaac	acaacagaag	atgtgaccag	ctccagagac	tcagcagagt	aagaaagtac	136320
attctacccc	cacatcagct	gcccacagt	caaactgaag	atgctctcag	cactgttctt	136380
cagcgcgcag	ctgtttaaagc	cttgccgtac	ctggcgaggat	caaggtcagc	acttagattc	136440
taattgtctt	ctctggctct	cattgttcac	tgggtttgtt	tggttttaag	gggttaagag	136500
ggctttcagg	ccaccatctc	agaagacata	atgctctgaa	gaagtatagg	aaccactcca	136560
ctcaatagca	aaggcttggc	aaaacagctt	ggcagcagtc	cacatgctga	cagtgtccag	136620
ctctggttta	ccagagcccg	agcacagata	acccctgagc	tgagttgcac	aaaacctacc	136680

agccacgaag	ctttagggcc	tctgggatgt	caggattcac	aatgacagt	ctggatgaga	136740
ggaccgagag	gctccgtcca	cogaagtcag	agactcgggc	tcctttgatg	gccatcacgg	136800
gctgccgaga	gccgtcaaac	ttgtcagcct	gcaaatacaa	cgcacgcaca	catcactgac	136860
aagaagaatg	tacaggatgg	ttttatggag	aagagtcagt	caatgtctgt	ggactctaag	136920
acagacctcc	cactcgggaa	ggagcattgc	actacaagaa	gctgcaataa	ccgatcatct	136980
cacacagcga	aggtcttcaa	atacttactg	gatagcacaa	ctccctgagc	taaagcccca	137040
ccctcaggac	tgggtcagct	gcaaggactc	acatcttctc	cccacagagt	tgtggtcacc	137100
accttccctg	acatgtccat	caaataagata	tttctcttag	caacttctct	gttggttcgac	137160
ttcactgtga	ttttaatcga	atcttcatag	ctcttgacga	ttccaatgat	gtctgaaaca	137220
aagaacactt	tgtaagagct	cccatcaagg	ctcctcttta	aaacacggca	gggacgaaag	137280
gcaagcacgc	tagtcacgca	tcaccccaaa	cactggagac	aaaaatcacc	actctttgcc	137340
ctcaaccttg	gacgcacct	ctagtgcgtc	tttagccttg	ctctctaggt	caccgatccc	137400
tgtgaaatca	aactgaactg	tgggttaagt	atggccatct	tcacagggaa	ggacagaagt	137460
ctcattattg	aaggtcatct	catagtcatt	tttaacagcg	gagaactgtt	tgtagcgat	137520
cttcaggggc	ccctttgaga	agtaatacac	ctgcaaaaga	ggccaagtca	gggcaagcta	137580
ctcagtcct	caaagccgc	cccacaatgc	agccttccaa	actgtgtcta	tggctctcac	137640
agttcagcct	gtcagctctg	actcaacacc	gtcttcttac	cttggttact	tcaataaggg	137700
gaaagaactt	gtccacttgc	tcattgaaag	cagtagctct	gatttcaccc	tgcagaagca	137760
agggagagca	cattagaact	gcgctgctag	gccctcgctc	tcaacagcga	gagcaggcca	137820
ttttgacagt	tagacaccat	cctcggagca	acagtcacag	tgtaaagaga	ccctcacagg	137880
gctgcagaca	ttgcagaatc	ccacatatac	agttacattt	tcaagctatc	agctctggta	137940
aaacaaaagc	agctcagcca	cccgagccta	cagttgttaag	ttaacacata	aaacgagggg	138000
gcctcgagat	gatctggaag	acaaaggtgt	ttcctgtgcc	aacctggcca	ccctagttta	138060
atcccgaac	cctcgggatg	gagaagacaa	tcgccccacc	aagttattct	ccagcctttg	138120
cacacgtgtg	catgcataca	cacgaaaata	attcttacat	cttatcaaaa	cattttttag	138180
aggaatagct	taatatcccg	ataatgaaac	ctaataattct	gggtcccag	tgatgggctg	138240
gaagcccaca	ggaaggcgctg	ctgaactcca	atgattcagc	actcttcctc	acgcccgcaa	138300
cctaagtga	ggcttacttc	acagtcagga	gctgggtcca	aaaatgttcc	aacaagtcac	138360
gctttcaagc	tgactgcaact	gttacttttc	tgtagggcaa	tatgttacag	atggataaga	138420
aacttaacta	aatggctaaa	accttaccca	gacaagctgc	aggaaattag	tttctgatac	138480
tgaatctgcc	atcacagtta	agatatctgc	tgtgctacac	cgtgacaaga	ggaagaagag	138540
ggggagagag	ccccaagct	ttgccccttc	cggtagcatg	gcttatttta	cgtgtgtggg	138600
tgttttacca	gcattgtatgt	gtgagcatca	ctgagtgtct	gatgcctgag	gtgaccagga	138660
gaggtactgg	atcactgaca	cgggagtcac	agatgggtat	gtgctactct	gttggtgctg	138720
agaattgagc	ccaagtcctc	tcaaagaaca	gcgagtgtct	ttaaatgctg	agcagtctct	138780
ctggcccctg	cattttgtct	taagtaaagc	ctagttagta	taaatgatca	gaagccctcc	138840
ccgcagacta	gttaaagctg	agtagctgct	gtcctctctg	ctggaggcaa	acccgcctcg	138900
ctcggcaggt	atcaccccag	ggcttcaact	tgccccctag	caatgtaatt	agagcgctgc	138960
agtctctgca	gacggagact	atttacagtc	caccaatctg	tattgctaca	gacgcgtccc	139020
tttaggagca	ctttatctca	tgtctgacct	tgtgcccag	atagcatcaa	aggtctccaa	139080
cagaaggaaa	gccacaatga	ccgccatgtt	cctcggagct	gagccccacc	cgnnnnnnnn	139140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	139200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	139260
atctctatat	atgaatgagt	atcctgtagt	tatcttcaga	cacacagaa	gagggcatca	139320
aatcccattt	acaggtggtg	gtgagccatc	atgtaattgc	tggaatttaa	actcaggacc	139380
tctagaagag	caatcagctc	tcttaacccc	tgagccccct	cccttttttt	tttttttttt	139440
ggcagcttat	ttttaattgt	tttaaatcac	gtatatgttt	ctgtatgtag	ctatgtttac	139500
ataagcgag	gcactcatag	aggtcagaga	tgttaagatg	ctttggagct	ggacttaaac	139560
atggttgtga	gccatctgat	atgagacca	agtggaactg	ggctcctctg	accgccttgc	139620
tgcttccac	ctagaattac	agtatatgca	attagctgct	cagccacccc	ctcagtcctc	139680
cttagtactt	tcaacatagt	gtaaatcgga	gagcactcct	tggtgcaaag	attctgtgtg	139740
ttctatcatt	ctctacaccc	aggcttcatt	gggacgtgtg	atgccagcga	tgattcattg	139800
gcgtgaagta	tgagttagag	tcagacaatc	agtgctctct	gctctgagtg	ggtctccctg	139860
cccctgtgtc	tctgagacag	gggaggattc	tctgtttgga	tcacatagtg	catagagccc	139920
tgtgtgcaca	gcgctctcaa	ccatttgttg	ccatttcaaa	tacagtgact	tcagagtctt	139980
cettgaaata	acagatttct	ccattttgtt	tggtcccttc	tttcgctaca	aagacctgag	140040
gatgagatgt	ttttgaagga	actttctagt	agttactcgg	tggaagga	caatgatgct	140100
cccctcttct	acagagaaga	aggaaacagg	aaattccaag	aaaaggagta	cacagatgct	140160
gcagtgtgtg	actctaaggt	aacgtgcgtc	caaccagcag	tgaaacaagg	cggaagatag	140220
aggcaagtg	cgactcata	accagtcct	tggtttgctt	tggttagttg	ggttggttgg	140280
tttggttgtt	tggttgtttc	ttcttcttta	ttagtcggat	tggttttggg	aaaaatgtag	140340
tagaattcta	ttgttgattta	tattcacaaa	caaaagactg	tttatatgtc	ctgaattttg	140400
gcaagattct	tcccatattc	cttcagaccc	atagcagcag	caagcttagt	ggccctttgc	140460
caggtattct	cactgagctg	tacagcattt	gtctgaagaa	ctgatgaaca	ttttagctta	140520

tgctcctatg	ttagtaata	gctagggttt	tagctagcca	ccttgctagc	ttacaactag	140580
aatgcctctg	gctgagtttg	gtggcacttg	cctttagttc	cagcactcag	gaggcagaag	140640
cagggtgaatc	tctgagttca	aggcagcctg	gtctacataa	caagttccag	cctggccagg	140700
gatatgtagt	aagactctgt	ctcaataaag	taaacagcca	gaatgactta	aatattgtcta	140760
aaaaagaaag	aaagaaagaa	agaaagaaag	aaagaaagaa	agaaagaaag	aaagaaagaa	140820
agaaagaaga	aaatgagtgg	atttgccata	gcctagtcaa	actagatttt	cttggttata	140880
tgtaacata	ctactattaa	caacaacaaa	tactttacac	taaccatgac	aagctatatt	140940
tttaaatat	tattttatat	gtatggtgtg	ggaagctggt	gcagagtggc	agttggctac	141000
tgctggccac	cacacataca	taggcagtga	aggttctttt	gccaagacaa	gttaaccaat	141060
cagatgtgag	acacgcctct	cctaggccta	tgtaagcagc	accagttctg	ggctcagggt	141120
ctcttcgcct	ctacaatcaa	gctctcccaa	taaacgtgtg	cagaaggatc	ctgttgagc	141180
gtcgtttctt	ctggccagtt	gagcgcgcac	aagagtattg	gttttcacct	acatttatgt	141240
gaactgcatg	tgcaacttgg	accaggggag	cccagaagag	ggcatcata	cccttggaa	141300
tagagtcaca	ggttatgtat	gggttctagg	aatcaaacc	atgtcctctg	gaagaacagc	141360
cagtgttctg	gatttaactg	ctgagtcata	tctccagccc	caccatgatg	gatttaactg	141420
ctgagtcata	tctccagccc	caccaagctg	gatttaactg	ctgagtcata	tctccagccc	141480
caccacgctg	gatttaactg	ctgagtcata	tctccagccc	caccacgctg	gatttaactg	141540
ctgagtcata	tctccagccc	caccacgctg	gatttgggag	cagaactgag	gctttgaaca	141600
ctgtgttatt	ctaactcttg	cttccttgga	accctgagaa	aattccttct	attggctttt	141660
cagagatctg	tatgggctta	aaacaagaat	atgtccaact	cttgatctct	gattttatat	141720
attaaaagaa	tagagtggag	ctaggagttt	ggcacactcc	tatgttccaa	gcttcaggaa	141780
gggaagaagg	tcaggagttg	aaggctagtc	agtcttagct	ttgtgacaag	tttgagcta	141840
acgtgagctc	tatgagccc	tgcttaggag	aggagaaaga	aaggggtaaa	gtgaggtgtg	141900
gctccacgct	accacttgcc	taacacgcat	gacgcccag	gtttccgttc	ccagccctgg	141960
ggttcagagt	attgtgcaca	ccctgtatct	ttaaaagtca	atagtcaacc	cttggaatta	142020
actttccgaa	aaatattaaa	gctacatcat	ccactctaca	aacttgtaaa	agcccttctt	142080
tgtaatggtg	taagaaagta	tttggttcca	gttttatggc	ctttgcatcc	taaatgttac	142140
cccatggaaa	gtgtcttaat	gaaaacaggt	aaaaataaga	caggagggac	tgcaaatgat	142200
ctcaggctca	caggtgctta	ctgcccttgc	agagaccatg	gcaggcagtg	ggtctcccta	142260
ctgcagggaa	tccaaaacc	tcttttgccc	tctgcaggca	accacattaa	cacatgcaca	142320
catacacata	attttaaaaa	taaaaataat	cttttaaaa	gagctctaga	acgagtttga	142380
catcagtcta	ggctacacaa	gaccttgtct	caagaagaaa	gaaatgaagg	tttgctgtgg	142440
atggagaggg	agatgcactt	cctattccat	gaagctgcta	ttttggtgat	tatggtactt	142500
tgcaatttta	tagagagctg	ctattttctt	tcttttaaag	taatgcttgt	tgttttgact	142560
gtaaaagtaa	taaatgttac	tttgaaaaat	atagagaagt	ataaagagta	aaaaaaaaag	142620
tcataaccag	tgaagtaccg	ttaacatttc	tgcttctatc	cggccagcca	gagtttttcc	142680
tgtagagaatg	tgtttttctt	ttacaaaatt	gggataatgc	tgcaacttact	gttttgtagc	142740
ccactctttt	ccttcacgat	ttattgtacc	cattttctca	cagtattaaa	ttttcagctc	142800
caagtaattt	tccatagcta	actctgtgtt	ctgtttacaga	gaagaaatgt	acttaattta	142860
agatctaata	ttggcataat	atttggcatt	atgatgctat	aataagcatt	tttctgtata	142920
aatattttta	tgtacagcca	gtgttttgtt	tgggatgaat	tagcacaaga	gtaaagtgtg	142980
gggtcaagcc	tctgagtga	gctgaattgt	tctcaggaaa	ggactagtta	acatccattc	143040
tgaagaatg	tggaatgct	catttcttaa	gcaacactgg	ttattattac	atactattat	143100
tttgttagta	ttatacattt	atttaggggg	atagaaata	tgctcatata	gttcaatttc	143160
aagagtttga	aaaatatatg	gtgtggtttc	ttttcctgtc	ccctagttta	gtctcacttc	143220
cagccccata	atctaccttg	ttaaatatac	tgtatataag	gcatatgtag	gtgaataaaa	143280
aacagcctag	tatggtggcg	ttaacctcgg	gttacttgag	aactgctctt	gcagagaaca	143340
tgactttggc	tccgaaagcc	ttcctggaat	tccagctcca	agggatgcag	tgccctctggc	143400
atccttgggt	acgtcactca	tgtgcacaca	tacacatttg	gtttttaatc	ttaggaaactc	143460
caagtgggcc	aatgagatgg	ctccatgtat	aaaggcagtt	atgcaaaggt	ctggatgaca	143520
tgagttcagt	cttcagattc	tgcaagataa	caggagagga	ccaacccctg	cgagttgtcc	143580
tctgacctca	gtacacatgt	catggtacgt	gtgtagtatg	cacatgcaca	gaagtcccag	143640
cactcgggag	gcagaggcag	gaggatctct	gagttttagg	ccagcctggg	ctacaaaacg	143700
atttacagtt	atataaagaa	actctgtttt	gaaaaacaaa	acaggggttg	gggatttagc	143760
tcagtggtag	agcgcttgcc	tagcaagcgc	aaggccctgg	gttcagtcct	caactctgga	143820
agagagagag	agagagggag	agggaaaggg	agagggagag	ggaggagagg	gaaaggaagg	143880
aaggaaaggaa	ggaagggaag	aaggaaaggaa	gaaagaaaga	aagaaagaaa	gggaaagaaa	143940
gaaagaaaga	aagagagaaa	gaaagaaagg	gaaagaaaga	aaggggaagg	aagaaagaaa	144000
gaaagacaaa	gcaaagcaaa	actaaataaa	atacatacaa	tataattaa	tttaagactt	144060
gagggcatag	atcagtggtta	taatgcttac	ctagcatgcg	taaaactctt	ggcttctaaa	144120
cctagcaccc	taccgcaaaa	tatttgctct	gtcttgctaa	attataattgc	tagttgtcag	144180
actactgtgt	actttcacta	gcaacataat	gagaatgttc	actatcccac	tcctctgtca	144240
agtaatctgt	tcttggtttt	atttttcttt	gccaaattga	tggtggaaca	agtatttcag	144300
ctagccctaga	acacacagag	aactctcttt	atgaactcaa	gtttcttata	cttttatgat	144360

ctccagaggt	ttgtttttgt	gggtttatta	gtgttttttg	tttggttggg	tgtttgtttg	144420
tttggttggg	tggttgggtt	tgctttactt	tttcttatcc	atTTTTttat	ttctttatTT	144480
ttttgttttt	aattttaatg	actggttcca	tgtggccaag	gatagcctca	actttgtagc	144540
agaaactggc	tttgaacttc	tgggtcttcct	tcatctacct	cccaagtgat	gggattaagg	144600
cacgtgccac	cacatctaac	aatatctggg	tttctttatt	ggagtttgaa	agggattcct	144660
ccagcattac	tttgactctt	catagtttct	tccagagtta	ttacactttc	atttgttaca	144720
ttaagagttt	gatccagggc	tggagagatg	gctcagtgge	taagagcacc	aactgctctt	144780
ccagaggtcc	tgagttcaat	tcccagcaac	cacatggtgg	ctcacaatcg	tctgtaatgg	144840
gatctgatgc	cctcttctgg	tgtgtctgaa	gacagctaca	gtgtaatcat	ataaataaaa	144900
taaataattc	tttaaaaaaa	aagagtttga	tccattttaca	ctggacttct	tgaggcagca	144960
ggatcataaa	ttcaaggtga	gcctgggtga	actggcagaa	gtggcagaag	ctgtgtctca	145020
cactgctgta	ttcatttcct	cattgatctt	cagaggtttg	ctaacgggaa	gtaagtggaa	145080
cagaaggttc	agtattcttt	ttttcccaat	tctattcagt	cttttagtagt	agatccctca	145140
ttatctgaga	tgcagagtcc	cctttattcc	tgtgaccatc	tcgttgtttt	tcagggagtg	145200
tctcattcaa	ggcctaacac	tgaggacatt	tactgtgct	atgccaatcg	ctctgcagcg	145260
ctcttccatc	tgggtcagta	tgaagtgagt	attgaagaac	ctggtgtcct	gcctgtggct	145320
gcagtggaaa	atgagctcct	ctctgttctt	ctgcacacat	tgaatcaaac	tagcttgcaa	145380
acactgacat	ccaccagac	ccattctctc	ttctgactca	tgtcacctct	cataggtgac	145440
cacaaacaat	atgtagttga	caagaagtag	ctatgtcatt	gtccacagtg	catggatttg	145500
ttccaatagg	ccagcacttc	tgtgtccata	tcagctagat	gtgctgctga	tagtatttta	145560
gattccaaaa	tgtgtccaga	tattacctcc	ttcgtttgtt	tcttctaaat	aagccaggca	145620
caaagacttg	aaagatggct	tgatggctct	tccagaggct	gggtttgatt	cccagcccca	145680
acatagcagc	tcacaatagg	ctataacgtc	atttccaagg	ggtctgactt	cctgttctgg	145740
cctctacagg	cagaaagcac	agacatacat	gcaggcaaaa	cacataaaca	taattgaaag	145800
aagatattaa	ataatagccc	acgcttgagc	ttattcctct	gatgatacag	ctctcctgaa	145860
gtcatcatgg	gcagtgtaaa	agtaaagggtg	ccccgccctg	cccaggggca	tcagtgaggt	145920
agctgactgt	gagtggcttt	cttctcatcc	cctaactgct	gcaacatcat	caactgtgga	145980
gcatttatatc	cgtggatttt	aagttaggaa	atgacaaaga	ttagatctat	ggccaggcac	146040
tagtgcacgc	ctttaatccc	agcactcagg	aagcagaagt	aggtggatct	ctgtcagttt	146100
gagtttacag	agagtgtcta	ggcagccagg	gctatataga	gaaattctat	ctggaaagaa	146160
taaacaaatc	agatctgtat	ttcaggaaga	tgcctatgag	ctgtttgaca	tgtgtgatag	146220
aggctcctaa	ggacaggaaa	gtattccaca	tgtgtgcatt	ttacagaaat	tggttatata	146280
ctggagttaa	ggactgttgg	aatagttgaa	tgttcacact	cagtgttctt	caaatcaaat	146340
agaagaacaa	gaatctatct	gggcatgggtg	atgcacaaact	gtattcctaa	catgtagaag	146400
actgaggcat	gctatgtgtt	tgtggctaac	ctgggctaca	tagtcaatat	tggacagtca	146460
gagctgctac	taaaacctaa	aaaacaaaaa	tctaataatc	tgggatttta	tattttcctt	146520
tttttaaaaa	aagagtgggc	agaatgtctc	tgattttgtt	cagatggcca	cagaacctag	146580
aaaaactgct	gctgctgctg	ctgctgcata	gcacacagct	aatatttgac	tatacgtata	146640
taaattttgt	tgtatacttt	agctgtgctg	tcaactttgg	aaaaaaagta	tcccagttta	146700
tcatttttaa	ttggcactgt	acagaaatta	acagccatat	tagtctagac	acattaaact	146760
tcattttttc	atttatacag	aagcaatgta	ctgtattaaa	tattcagctc	tatctacagg	146820
ggtttgatta	cagaaactat	caaagtattc	tctaatagat	gaaaaaagat	tcaagaatct	146880
gactgtagat	ccaaaggaca	agtggagaaa	aaacttaggaa	gaattttccc	tttatcccc	146940
ccctatatatt	atcatctctt	ttacttctaa	taatagtggc	cattttattga	acataccagg	147000
gagttccttt	catcacttta	gatatataat	ttatctcatc	ctaaaatgac	ctgttgatga	147060
gtcatctctc	tttcagatga	gaaacaaaga	cattgaaaaa	tctaacttgc	ctgcataaga	147120
tcacacctag	cctcttactc	actccatgaa	tattcctttt	ttttttttcc	tttgagacag	147180
aatctcacta	tgtagttctg	gttgtcctag	aactcaatat	atagaccagg	ctagcctcaa	147240
actcacagag	atctgatagc	ctctgcctgc	cgagtgtctg	ggttaaattg	atgtgtcacc	147300
aagcccagca	aaatttacct	tcttaatttt	ctaagacgtt	tctcctctta	aaaaatggaa	147360
ctattgagcc	agtcatgggtg	agacaggctt	aatctttaat	ctcagcactt	aggaggcaaa	147420
gacaggccta	tgggttcggg	gcagcctgat	ctatagagag	agttctatgg	gttaaagttt	147480
agggttaaag	ttttgagaca	aaactttgtg	tcaaaaacaa	acaaacaaag	ccagactgct	147540
taataagaca	aatcagacat	aatattataa	acaagtatta	gtgtcactca	attaaaaagt	147600
cactcaggag	gtgagatcaa	tccccagcat	aatgagggga	ggagggggaga	gaagggaatg	147660
aatgggaggg	gaaggagga	gaaggagaa	aaagggaaga	tctcaaagca	gaacacagga	147720
tgtaatTTaa	ggcctaagct	ctcgactgaa	gttgtccact	tttaatgacc	cttttcatgc	147780
tcattggtttt	ctgtcttcgg	tacatagtga	agagtggaaa	ccaaggggtca	tcctggaatt	147840
tcctttttgt	ttcaggcatg	tcttaaagac	atagtggaa	caggtagtca	tgggtatcct	147900
gaaagactgg	agcccaagat	gatgggtgct	aagacagaat	gcctgggtgaa	cctggggaga	147960
ctccagagtc	caagacagac	catcagtgat	ctcgaaagca	gcctcactgc	caagccaacc	148020
ctggtgcttt	cctcttacca	gattctgcaa	aggaatgtcc	agcatctgaa	aataaagatc	148080
caagaaaagg	agactctccc	agaacccatc	cctgcagctc	tcaccaatgc	cttcgaggat	148140
atagccctgg	gggaagagaa	cacacagatt	tctggggcct	ccctctctgt	cagcttatgc	148200

acacaccctt	tgaaggccg	ccatctagtt	gccacaaaag	acattctccc	aggagaactg	148260
ctggtgaagg	aagatgcttt	tgtaagtgtc	cttatcccag	gagaaatgcc	acgacctcat	148320
cattgccttg	agaacaagtg	ggataccaga	gttaccagtg	gagacctcta	ctgtcaccga	148380
tgtctgaagc	acactttggc	cacagtacct	tgtggcagct	gcagctatgc	caagtattgc	148440
agccagggaat	gtatgcagca	ggcatgggac	ctctaccata	gcacagagtg	ttctcttggg	148500
gggctgctcc	tcacactcgg	ggctctctgc	catgttgccc	tgagaatgac	tcttttagcc	148560
agatttgaag	atgttgatag	agttgtaagg	atgctttgtg	acgagggttg	tagcacagac	148620
acctgtttac	ctgaaagcaa	gaatctggtc	aaggcatttg	attacacaag	tcagggagag	148680
agtgaagaga	agagcaagat	aggtgaacct	ccaattcctg	gatgcaatgt	caatggaaag	148740
tatggaagta	attataatgc	tatcttcagc	cttttgcccc	atactgaaaa	gcatagccca	148800
gaacacagat	tcatactgtg	catcagtgtc	tccgcactgt	gcagacaaact	caaagctgac	148860
agcgtgcagg	cccaaaccct	aaagtcccct	aagctgaaag	cagtgaacccc	agggctgtgt	148920
gcagatttga	ctgtttgggg	agcagccatg	ctgcgacaca	tgctacagct	gcagtgtaat	148980
gcccaggcaa	taacatccat	atgtcacaca	ggttaagtcag	aaatgggttt	tacttacatt	149040
attggatttt	caagagctaa	tgtttaagga	gaaaaacact	ataaaggaag	cctggcatca	149100
aataaatcag	tgacctaaaa	ggaaaacaca	gccgtcttat	atatcattat	gctatttgaga	149160
agctttgagc	acatttctgt	gaacccagag	cttgggaggt	ggagatagga	tgattaggag	149220
tctaagacag	ctttagctat	acagcacgtt	tgaggctcagc	ctgaactaca	tgagaacttg	149280
tctcttaaaa	acttgagcca	gagccagggt	gtggtggcac	atgcatttaa	ttctagtact	149340
caagaggcaa	aggcaggcag	atccctgaat	ccagcctcgt	ctatatagtg	agatccccac	149400
caggctacat	agtaaatcc	tgtttcaaat	aaataaatat	aacaaaaaca	gcaataataa	149460
caatagcaac	aaattaat	tttagatgta	tttatttatt	ttatgtatga	gtacaccatt	149520
gcttttttca	gacacaccag	aagaggcat	tggatcccat	tacagatggg	tgtgagccac	149580
catgtatgtg	gttgctggga	attgaactca	acacctctgg	aagagcagtc	gggtgctctta	149640
gccactgagc	catctctcca	gtccattaat	taaaaattta	aaactagagt	atttttaaac	149700
atttattcat	tttgtgtgtg	gtatacatat	tataatacag	gttcataagt	caattctctt	149760
ctaccatgtg	tgtcttggag	atcaaaactca	ggttcttagg	catgggagca	agtattttact	149820
tcttgaacca	tctccctagc	catttctagt	attcttttct	tttgtcttga	aagattttatt	149880
tattatatgt	aagtacactg	tagctgcctt	caaataccgg	aaaggggaat	caggtcttgt	149940
tagagatgat	tgtgagtcac	catgtgggtg	ctgggatttg	aaactcaggcc	ctccagaaga	150000
gcagtcagtg	ctcttaactg	ctgagctatc	tccagcccca	tttttagtat	tcttattaag	150060
tggtttccat	tttatccaaa	gatgccttta	agggcctggg	aagatggctc	agtgggcatt	150120
gaacttgggt	tgtgagcatg	aagaccagag	ttcagatccc	tagcaccacag	gcagatgctg	150180
aatgatgggt	gcctgcctga	gattccagga	caacggagac	agacaggggc	cctagctaac	150240
catactacac	actagctgag	ctgtgtgctc	aagagagcag	ccctggctta	ctgtgcagga	150300
tgagagagtga	tcactctccac	aggcaagcac	acacctgagc	acacagacat	gcacaaaagga	150360
aagaaaagtc	cttttaaggt	gggtggtggt	ttgttgtttg	ggggtctttt	gttttgtttt	150420
ttctccccc	tccctcattg	tgatggcaca	ttcctttaat	ctcacatctg	ggacaaaagag	150480
gccggaggat	ctttgtgaac	tgagggtcag	cctgttctac	atagcaagcc	catttcagcc	150540
aggacgacat	agatataccc	tgtctcaaac	agacaaaaat	tatttatttt	atatatttga	150600
atgttttgcc	tgcatgtatg	tctatgcaca	ttatgtctgg	tgcccatgaa	agccagaaga	150660
gggcatcaga	tctctcagaa	ctggaatttc	agacacttat	caagtactgc	ctgagtgtca	150720
ggaatcaaac	caaggtcttc	tggaagagca	gcaagttagtc	tttttttttt	ttaatttttt	150780
tttattacat	attttcctca	caatgtcttc	tcattgtctg	ccaaaagtcc	cccataccct	150840
ccccccccc	ccccgagcag	caagtattct	tcattgtctg	gccatctccc	catctccttt	150900
tctagttaat	taagctgaaa	gggaggagg	tagatgttgc	ccaaacttag	gatttattga	150960
cagattaata	ctctgttagc	ctaactacac	tatagaagct	tattcttttag	actttcacat	151020
tacactgtcc	agattttgcc	atcctttttg	ngtgtatatg	tctacagatc	ttaattcagc	151080
tgccaattta	tacagtgttt	ataggatttc	tttctgacgt	ggatctttta	cccatcttaa	151140
agcagtagga	tttgaaagct	gacatttatg	tggcctatgg	tctgtttaa	tcacatttca	151200
agttagtctc	tgtggtacac	attttgggg	ctatctgcgg	ttccgcatct	cacacttttc	151260
cctctcaggg	tgtccagaag	ctgctgcaca	ctgggctgga	aggatgaagt	ggagtcagga	151320
gtgagtggaa	ttctgcagca	tcccggtcca	gctgggagtg	aatgctgggg	tcaggaggag	151380
atgggtgaga	gggcttctc	caagggcctt	cttagtggtta	cagctctagg	caaaggcctt	151440
ctctgacaat	cttagcctgt	gcatagtttt	ttatctcgaga	tgagcttgta	tgcatacact	151500
ttattggcag	taaatcagag	gttatccact	cttaggggaag	gagataggaa	tacccaaggt	151560
ggacagaggt	cattggctga	aggataacgt	actgagatgc	tcattagcac	ggggaggcat	151620
cccaggaatc	tcaggtgctt	gctgactggg	tttcttaggg	ggttgagagg	gtagcagtga	151680
tttcaccaag	gttatgtatg	gcagagggtg	taggggtttc	aggctcccca	gacaaaagaag	151740
gagaaggaga	agccctgctg	ataagggaag	tcccctattt	tgagccactt	cagaaggcta	151800
tcaagacact	gatagacttt	gtccttaatt	agcaggggcc	aacaagtgtc	tgttttcttt	151860
tctgccttca	ttggctcttt	gagccactgc	tacaaaatat	ccaaatctgg	gccagggaagc	151920
tggtcgggt	agtaaaaggtg	tttgctctca	agcctgaagg	cctgagtttt	cacttgattt	151980
ggtttggttt	ttgttttttt	gagacaaggt	ttctcagcat	agccctgggt	attctggaac	152040

tcactctgta	gatcaggctc	aacttgaatt	cagagatctg	cctgcttcta	catccccgagt	152100
gcttagatta	aagtgtgctg	ccaacactgc	ccacctaaaa	aaaatatgag	gggctggtga	152160
gatggctcag	tgggtaagag	caccgcactg	ctcttccgaa	ggcgcgaagt	tcaaatccca	152220
gcaaccacat	ggtggctcac	aaccacctgt	gatgagatct	gatgccctct	tctggtgcat	152280
ctgaagacag	ctacgggtgt	acttacatat	aataataaat	aaatcttaaa	aaaaaaaaaa	152340
aagaacacta	ttcacaaggt	acaacaacag	tgtactagga	aaactttaaaa	tagcccatca	152400
ttcattctag	gggaaaaatc	tttttaaact	ttagtgtgta	agaaagagag	aggggctgta	152460
gaaaggccac	agcagtgta	tccagggttag	gagtcaactt	ttcagaagcg	gagtcctccc	152520
ttctacctgt	ttttgaggca	gtctcttgtt	tctgcctac	actttgtaca	tgaacttcaa	152580
gatgggtgtt	ctatttctgc	ctctcatgtt	gccctatgca	tgctgagctt	actgatgcc	152640
gccaccacat	aagcatggca	ccagcactga	ggcatctggc	tcacatggag	152700	
agttacccat	caagccatct	tgctagcccc	agaaaatgta	tttttgacag	gtgtggtggt	152760
gcacatattt	aatcccagca	ctcaggaggc	agaggcaggc	agatctctgt	gtctgaagcc	152820
agcccgagttt	acaaatcaag	tcccagaata	gccaaaggcta	catagagaaa	ccctgttttg	152880
aaaaacaaac	atccttctgt	gttccagtca	caatgactgc	tgtaaacaata	atatgaggat	152940
ttgggcgtgt	caaaaatcac	aagtgcgaag	gtgtaatggg	caggccttta	gtctcagcac	153000
ttgggagcca	gaggcaggag	gatctctgtg	agtttagcac	agccagggct	gttacacaga	153060
gaaaccctgt	ctcaaaaaaa	ccaagcaaaa	atagaattac	aagttaacca	gggtattggt	153120
gttaatatga	aatagcagaa	ctcaagatag	ctaataaaac	aaaggattct	attttataaa	153180
ggagacattc	catactgaaa	tatatgcaga	gcctgatgct	tgccatgcaa	ctgaactaca	153240
cnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	153300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	ngcctgggca	agagtaagca	153360
aaccgtcagg	ctcagaggca	ggatggcagt	gatgtgtttt	ggggcacaag	gccccctttc	153420
attaaagcac	caaatcttgt	actaaaaaca	gccctctgct	ctcgatctgc	agagatacga	153480
gagaggcaca	catagcttcc	gggcccttgc	cctctgctct	cccggctcag	tcctggactc	153540
ctgggagaag	cttgaagctc	aagaactggc	cgctgttgcc	tttgtgtgta	cgccagaacc	153600
agcagctatt	cagcagctgc	ggctcttggg	caagcttggg	agtagctcgt	tccttctcct	153660
tctgcagggg	aaagacaagg	agatggcact	gagctgggta	gccaggatgt	gggaagtaaa	153720
ttgtgtctat	gtgggtgagg	gaagtgccgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	153780
gtgtgtgtgt	gtgtgacaga	gacagaaagg	gagagtgcac	gttcttgtgt	gttctctgtg	153840
gtcagtgtct	cctttggctc	ctcctgccaa	aaagcatgct	gtgttctcag	agtcaagctt	153900
ggccaggctc	gcccacccca	ccaaagccca	ggacttgcca	cccattaaag	aatgcggggg	153960
taggaaataa	aaaaacagggt	tcttgtccaa	gagaattttt	attattattt	tttctctctt	154020
aatgggattc	agcatttctt	actcctcagt	atcctctctg	gtagggaatt	caggctctgtc	154080
tatgcagagc	acaagagact	gtgcttgcca	gaaacccttg	ggccaacagc	cctattccct	154140
gactgggctt	gcctgcaggc	tgcttctctg	gctgacccct	gagtcctggc	ctctgacctc	154200
tgcccgctct	ggggtcatcc	aggaggagca	ggaccactgt	gatacagggt	tcctgagatg	154260
gctactgaaa	accctcatc	tgctaaggcc	actgattttt	tttccactg	ccagactact	154320
caggggccctc	agtaggtcac	tggccaaaga	gcctagatgt	taaaactgca	gctaaggcct	154380
ctcctgaggc	cagagctcag	agcctccctg	gcctgcacaa	agtgcctaaga	aggacattgt	154440
ccatccaacc	catttgacta	acactgtaga	cgctgccttg	tctgccagct	tgaaaacccc	154500
agaaacctct	tcccagctc	cgctagctt	gcctccagca	ccctgacatc	cacttctctc	154560
gctaattgtct	gcagcttcta	cagtagggtg	gacgggggtg	tccgggggtg	cagacaggcg	154620
tgtgttgcat	ataaaaaaga	ggtgatgttc	taagtattcta	agaatgttgg	tccctggaag	154680
tgattcttgc	tgcttctctt	ccttcccagc	tcttcccact	gacacctttg	ccccagcaa	154740
gcccaggagc	tattgttgct	ggctggggtt	ctgaacaaaa	ttgccgcagt	tttttgtttt	154800
tttttttctt	gctagaagtt	accgatacag	tccttaattg	agcaaatata	ggttcccgtg	154860
tagtttataa	acatgataag	acatacagtt	tggttaaggag	tggttgagg	gacctgtgca	154920
tttatatat	tatatatata	tatatattat	tggtgtgtgt	gggcagagt	aggatatata	154980
taagtggaca	taggtataaa	actgcacctt	ctgtgtgact	ctcattgcca	gtacagttct	155040
aaatgtcatc	cactggcgat	ctctcctttg	gtgattggtt	cttggaaccc	agcaggctct	155100
ccgggggctg	cctgagacgt	cagggaatgag	aggcattacc	cccatggata	gggactgagg	155160
gtggcatagg	gttgacagg	gcaggttaac	taagtgatct	cagacaagag	ccaagagtgc	155220
tctgagattg	ctggttgccc	cagctggctc	tgggagagcc	ttgttctgag	tcctgtcctt	155280
tccaaaacca	gcagggtcct	tcagcccttc	tctccaaatg	accaggcttc	cgagagcccc	155340
agcttcttca	aggggcgcat	gtcccagcac	cactaatgac	tacttttgcg	tgcttttgac	155400
cactgtgctg	gagtggatac	ggtccagagg	cgctcggtca	ggacagccga	gtgagacgtg	155460
atacccttcc	cgtctacggc	tgtacatttt	gggcttataa	tccaccagga	agggtcagga	155520
cggtggctga	aggcttcagc	agccttttcc	tgaacccag	cagatcttcc	acttaggaaa	155580
aaaaaaagaa	agaaagaaag	aaaaattctg	ttccttctg	gacatgtggc	155640	
agtgtgggtg	gacggagccc	tgggtcctca	gctggagagt	actgccagcc	ccagtatcca	155700
ggccaggagt	ggggcccaaa	gggcccgtgc	tgaggatgcc	tgctgcaccc	cactgggggc	155760
acggatgggg	gtcctgcgag	cacacttgcc	cttctctctg	ggtcgtgcgg	tgggcatgat	155820
gtcaaaagctg	aacttgtgct	gatagtcggg	ggcatagtcc	tgcatgtcgg	taagctcttt	155880

cccagagctc	accttagaga	tctggttccg	gttcctgtgg	ctggtgcagt	tcttgccctgc	155940
cttcttgtaa	cctgacctgg	agccaggcgg	atggccatgt	gggtggcctt	tgtccctgga	156000
ggccccatgg	gacggatggg	gctccttgcg	ggcagccctg	tcagagggtg	taagcgtgtg	156060
agacttgatc	tgggtgaggag	acactgggtcc	tgtgcagttc	cggaaagtcct	ccaccctcag	156120
cagcttcaga	tcctggcctt	gccgcagctc	gggggtcgcg	caggggacag	cagagctaga	156180
gccacggaac	cttcgcagcc	attcccacag	ggaacgtgcc	cggcagccac	agtcccaagc	156240
attcccattg	aggcgaagga	actccaaggc	caccaggggg	gccagacagt	caccctgcag	156300
ctcagtgagg	ctgttggtga	agagaaagag	ggtggttagc	ctgtggaggt	catggaaagc	156360
cttggtggtga	acccactgta	gctggttctc	atgcagcagc	aaccggtcca	ggttcaccag	156420
gccccggaag	atgccttggc	ccaggctcca	tagcttggtta	ccatggagaa	acaagtgact	156480
gagattgacc	agggtccacaa	agatgtcatc	ttggagggtac	tcgatatggt	tgtcctgcaa	156540
gtagagatac	tgcaggctgt	gcaggccacc	aaagatgcct	gcgggcaggg	cgctcagtcc	156600
acacttatag	aggtagaggg	cgtgaagcct	caccaggcct	tggaaaggtct	cgggtgccag	156660
cgttcgcagc	tgtcgggtgt	ctccaaggte	tagctcctcc	agatgcacaa	agccctcgaa	156720
ggtgttggtga	gcaatgaaag	tgatgttggt	ggagtagatc	cagaggggtga	ccatggcggg	156780
gctgaagtgg	ccctgctgga	ggaagggtgat	gcgattgttc	tgcaggaaga	tgcgctcact	156840
gtcctctggg	atgccctccg	ggatggcagc	aaagtgtgtg	gcctggcagc	tgacagtcat	156900
gggctcgagg	tagcacacac	agtctcgagg	acaaccacca	cccagaggtta	gctctccagc	156960
gagcagcaac	agcagcaatt	ccacacagca	ccctggtggg	gagagacaga	acagcagtga	157020
ggggctgccc	agaggaggtg	gagatagatg	gggaacagag	ggtggaatgg	gggactggca	157080
aatgactctg	ttggctcaca	gagggttctgt	cctctgtatt	gtatgcaggg	gtcccttgga	157140
cagggcattt	ggggccaagg	cccacattat	ctcctcacct	ctttagctct	gtccctaagg	157200
tctctaatte	cattccgaaca	cttcttcaga	ctgtgcagcc	caccgcaggg	tggaaacacac	157260
ttgtgaacac	aggcagtgctg	gcctccggct	ctgggtccgg	ctctgccact	cgctcactgt	157320
tagctgcctt	agcaaggaat	gactctaaca	aagcaaactc	ggagtccctga	atgatcactt	157380
tatttaaata	attcctcaaa	ataaagaaag	cattgagtcc	atggtacca	agcatgcctc	157440
aataagcgcc	tctttcacac	tgtggtacaa	aaacttcaaa	cctacaactc	ctccatggct	157500
gtttcctcat	gagttaaaca	cagttcacag	ggctgtgtgt	acagacaagg	cacatttctg	157560
tgaggggctg	tgagtgacac	cagggctgac	gcacgaggct	tcocctgggg	ttcacagtac	157620
tgccagatcg	aggctgcatg	ctcctctccc	ccattcacac	ccccccctcc	tgtgctggag	157680
attgccaggc	tgtggctgta	aaaccggggc	ttgcctcttg	actgtccaga	gcatttctct	157740
tgtagcttcc	ctctaattgg	gcattaatta	ggcattcggt	aatggatcct	taaaataatt	157800
atthttcgat	gtgtccagcc	tgtggtcggg	taataggcct	atgctcatta	tggaaagccgc	157860
ctcattatgg	acgattgtca	ttacctgcct	ttttccaggg	tcacagctgc	cccaagtggc	157920
ccagcaggcg	cgtcaggaag	atggggacag	gctccaggcc	tacgggcgcc	caccctgaag	157980
ggccaggcag	ccacgacctta	tgtcgctcca	gttggcctct	tgcccttctt	tttccagctt	158040
gttcagctgg	gactcctggg	agagccaggg	ccctgggggg	aatatgagct	gagctgaatc	158100
ttcttgctgc	tagctgtgct	cagagcaagt	ggaaggagca	gggaccttct	gaccaggctt	158160
cccacttggg	gtcccaggcc	cagggactgc	gcaggccccg	gcagagtagg	ttgccacctt	158220
gactcttgac	gccccccccc	cattcccac	agaaacagca	tcgtaagttg	acagctccca	158280
gctgttggtga	gttatgggct	cccagagagt	ggcagctgct	tctcgtccct	gtaatcacc	158340
ggcttcagct	agaatgtttc	tagcacataa	aatcatcgc	atataattta	gtttttgcat	158400
aattgggttc	agttgtgatt	tcagagcaat	tatgacctca	gcagcagggg	tggcacagca	158460
taggacccct	ttctggggccg	ggccatgccc	tgcaggccct	ctgcagtgtc	ttctgcccac	158520
cggcctttag	acagatgca	ggctataacc	gtctctgctt	caatttctctg	ctccaaaaagc	158580
tgtctagatg	ttactctgtc	gatcttctct	ctcagcatcc	tgggtgtggc	ctccagcctg	158640
ccagcctctg	tcctagggat	cctgtgctgc	agagggaggc	acagtcggag	ggaggggagc	158700
ctgcctctgtg	ccaccagcac	tcacactggc	tgacagctcc	acagaccac	agctccaacc	158760
tcctgtcctt	gcttgacccc	tctcttccag	gaaggccatt	cttgccagaa	cctttcccaa	158820
cggtcctcctg	ggaaagcctg	gactctaggt	tcaaggacat	tcagtatgct	tgccccacat	158880
tttatgctgg	atgagacaca	gcagagcctt	cttactggg	gggtcctgtg	aaaatgaaag	158940
cttttcttcc	ccggcctgca	gctgcaggca	ggtagggggt	gcagtgggct	tatcactaat	159000
accattcgac	atthttgacag	ctcatcgagg	tttacagagg	gcttttggtg	taccctcaac	159060
ttccctctgtg	ttctcccacc	tactgtggct	gctctgtctc	tgtgcacccc	aaaagaatct	159120
ggaagtccct	tggggagatt	accocctta	cataggggcc	tcaaaggaat	acagggtaca	159180
gctctattta	ggaaaaaaaa	aatcaaaactg	aaccaaactc	caggtgtgga	cttagtaacc	159240
agttttataaa	cataccgtgg	caagtggagg	aggcaggcgg	cagaacggca	tgaaggtagg	159300
actgggggttt	tcctcttaaa	agggcaatgg	ggggacaaa	ggactctagc	caagacctga	159360
tgcagaacca	ggctcagttc	ccctgttatc	tcaaggctat	catcactagg	aggcctaagg	159420
caatggacca	caaggacctt	gtccttgtag	gcagctcact	tcctgcagtg	aagtgcctct	159480
ctggaagcat	actaatagga	tgtaggctca	ggacagctgg	tctctgtcct	ttagtatattt	159540
tccacatgcc	agggatgtta	accttccaag	cctccatctc	ttctaattggg	gggggggtgt	159600
tgaggggctc	agccactctg	catccatgtc	tttgaaagcc	agtgggtatta	ctccaggacc	159660
ctgagcaagc	tgtcctagtc	agccctggcc	acttctggac	tccttgcttg	agtcagtagg	159720

tgccaatcct	aggattgtca	ccagcaggtt	tcttcctagg	gaggcaagca	ctgtatcacc	159780
atggcgccct	ctatgcccc	tctatgagge	ccttgggagc	cccgccccc	tgattgcctg	159840
attaatgtac	caacaatgag	gatggagcct	ttgccatgca	ttttaacatt	gcaaattagc	159900
aggaatccaa	gtctctgtgg	aggggcccctg	cacctcttct	gccagactca	tcaagcgctc	159960
cttgggcagg	gctgccttct	acttgagggg	gcggaaagga	gaagaccag	ttccactctc	160020
cttccccctcc	aggaggtgcc	cttcatcgctg	ttctgtctcg	ttactctcaa	gcctccggcc	160080
tcccacgcac	gtgagctccc	aaggggctct	acagcctccg	tcattccttc	ttccattcat	160140
acttgcctcc	tagtctagga	gagccatgga	agacagtgtg	ggaagggtct	gacaatgagc	160200
atcatgcccc	atttgcatat	gcggtggcaa	taccctgggtg	ggtaccagga	gagtataggg	160260
gaaattaaga	gaggggccta	aggaagcct	ctgctatccc	tgggctacca	gtcagcattg	160320
cttgggtcact	gatccccctct	gtaacaccag	cccttctgca	acctgccaga	gtttttgacc	160380
tttgaactag	ggctgagaag	ggtctgtctct	gttcagctgc	cttggctggg	aggggaatct	160440
gctcagacct	cagcacacac	tcaacagaag	gcatgcaagc	aagggaagcta	gcagtggcct	160500
tgggtcagct	ggcaagcccc	aaactcttcc	tgccaagctg	agcatgaaaa	gccacctcac	160560
catgggtccca	tgggaccaga	cctggtagga	taggtggcaa	ggctaaggca	gcggaatagc	160620
atgtgcaaag	gcaactgggg	gggaaagggc	ctgtgcttct	cacccccctct	aatggtgcag	160680
agcctccaag	gaatactgta	acctcagctc	agctgggctc	gggtggccag	agagcttggc	160740
accagaacca	gcatcaacag	ggcctgtctg	ctaaacccag	acctcacaag	ccagttagt	160800
aggggcccctg	tagcaccctg	gccaccagaa	ctaacgagga	agatctgacg	ctgggaatat	160860
gtctttaatg	aaaagccctt	ccggaagcca	catttgcaca	gaagaaaatg	aggtgcccag	160920
agcatcagtg	ggctggttgc	agctggagaa	cacagcaggg	ggacaggctc	taccaagcta	160980
ccctggccttc	aggtctgggg	tctagccagc	tcctgtatgc	ctggagtagg	taaagcagcc	161040
tcgaaatggg	ctgggtcagc	tttttcaggc	tccaaaggg	caggacagct	gctgcagctt	161100
agcacccaag	ggggctgccc	ctctacccct	aagtagaggc	atccccatgg	ccccctgggca	161160
ggtcagtggtg	tctctctgaa	gctttgtagg	ctgcttcttg	gccatgtagc	caatctctct	161220
ggccttcag	ctccccctcc	tgcctccagc	ctggccagct	gctcttctct	gagcaatcga	161280
tgtaaccga	atgtctctct	gctgtgggga	tggcgccctc	aggccaggcc	agctgcactc	161340
ctggggctgc	tggcgccctc	ggccacttgg	cacttgtgcc	acttgtgttc	taaacacagg	161400
ctccttctctg	gctcgccctc	gaagacaaga	aagctggcca	gggaacagct	gggctcccat	161460
ctcagcctcc	actgctgtgc	agagcgccg	gcagcctcct	atccatggct	gtgagtagaa	161520
cagggtctgtg	gagccagagg	cctaagttga	atcctggctg	ctccttttaa	tgcttgccag	161580
agcctcggtt	tcctcacctg	caaaatgggg	cagtctattg	aagctcagcc	agtcctccag	161640
ccacagata	ctggtaccca	cctgccttcc	ccacatctcc	atctgtctaa	ctgaaacat	161700
ccaaaccaag	ctcttctctt	cctctccggc	ctcctccgtt	cacaacttct	ccatcttggg	161760
taaaagtgg	ccctttacat	cagctgcctg	ggaggaacac	ctcagaatca	ccgtggctca	161820
ctctggggaa	attctgctgg	ctagatttta	gaatgtatcc	agtatctatc	cacttatcat	161880
actcgctatt	gtaccatttc	accagtagc	ctcctggatg	ccctcccccc	ccccccgcaa	161940
gccttgccctc	ctcaccccta	cacctccttc	aacagggaact	agggtagtcc	agggaagtg	162000
agtcaggaag	ggctgctcct	tagtctgcat	cctccagagt	gcccatactaa	ctagaagctg	162060
ccccaggtct	ttctcccagc	ccagaggccc	tgctctctcc	tggctacaca	gacctgtctg	162120
ccgtccctcc	cacatgccag	ccctcagcct	ccctcctgcc	ttgtccatg	ctgttccatc	162180
tacctggacc	ggttcccagt	gtgtctgcag	ggctgtcttc	cagagaagcc	actcttgagc	162240
agcgatttgc	aaggagcttc	tttctctggg	cattttactc	tacagcacgc	gacctcttgg	162300
acagcacgac	ataagtcact	tgtttctctt	atgtccgctg	ccactttgtt	ctgatgagtg	162360
tgctccctgc	acacagtagg	tgctcattaa	cagttctggg	ggagggaatt	accttctcaa	162420
gtctctggag	aattgaatga	tgacactcag	gaagccctag	gctcaagcct	ggggcctgga	162480
tcatagtagg	tgctcgataa	atgttggttg	taattagtcc	tgggagactc	agagccttca	162540
ggagaacaga	cacctgaact	tggtcacat	caagactcct	taggcccata	aaggagtga	162600
ccgtttgttg	gatgagcact	ctgagaagga	cccaatacca	gtcctttcct	gggcaagggg	162660
aaatagactg	ggactgggga	gtttcccaag	gtatgggatt	ttcagcacta	aactggaata	162720
atgctaaga	aaaaaaaaag	ttagtcactc	taaaaggggc	caggaccaaa	acctttcaaa	162780
cagaaatgtc	tgggtttatg	aagagaggaa	gccagatatg	gtggggcaca	tctttaatcc	162840
aggtgcttgg	gaggcaaaag	cagatggagc	tctgtgagtt	tgaggccagc	ctattctaca	162900
aagtgaagtc	taggacagcc	aaggctacat	agagaaaacc	acttgacttg	ccaccacaa	162960
taaaaaatctt	agtgggggag	cagtcaagga	ggaaacacac	acacacacac	acacacacac	163020
acacacacac	acgttagtat	aatatcatat	tatggctctg	tgcttgcagt	ccagggaatga	163080
gggtcgaact	cagagtgtta	gtgtgtgcta	gtggatattt	gagctctgta	tttatgtgca	163140
tgctctgtga	gatgtgtacc	tgaggtgttt	atgtgtacac	aggtgttggc	ctgttgcata	163200
tgagtgagga	catggttgtg	tttctagagc	atgtgcatgt	gtctgagttc	atgcacataa	163260
actcacatct	acctctggag	actgagagtg	acaacccagg	gcccttttat	cctgcagcac	163320
cccaggccca	gcaccccgac	ccagcatccc	aggcccagca	ctccagaccc	agcaccacag	163380
gcccagcatc	ccacgcccag	catcccaggc	ccagcacccc	agaccagca	tcccaggccc	163440
agcaccacag	accacgaccc	ccaggcccag	catcccaggc	ccagcacccc	gaccagcat	163500
cccaggccca	gcaccccgag	tccaagcact	caatgcccag	caccccgacc	cagcatccca	163560

ggcccagcac	cccaggccca	gcattccaagg	cccagcatcc	cagggacagc	accccaggcc	163620
cagcatccca	ggcccagcat	cccagggaca	gcaccccagg	cccagcatcc	caggtccagc	163680
atcccaggga	cagcacccca	ggcccagtat	cccagggaca	gcaccccagg	cccagtatcc	163740
cagggacagc	accccaggcc	cagtatccca	gggacagcac	cccaggcaca	gtatcccaca	163800
tggaggcagc	acatactgaa	gatagggaat	gtctctgagg	cctcttatct	tggtcettac	163860
cctcattgct	ttcagcacct	gctctcctca	cactcggaat	caaacaccct	gtgcaggttc	163920
tcccagtacc	aggattcccc	tcagctgagg	aatgggtagc	taccattttg	gcttttgtct	163980
gtctggggtt	ggcagcccca	tgctaattgg	actgacagtt	tctcctgaga	gcaatttggg	164040
cagcacatcc	tgcccattag	gcctaacctt	gcctgcaggg	gtgtgctgtg	ggggcaggga	164100
tggagcctac	cctgtatagc	tctgtattga	ggcactcccc	caagctatga	cccatgccag	164160
tgggagtcac	ttcaccttag	caactccaga	tgggcacaaa	aatctctcca	ataagggtag	164220
gtatgggaat	aggtaaaggag	agcatagtga	gcctggctgg	gcacctgaga	cctgagcagc	164280
ctgcacggga	gatttgttca	ctgtggttcc	agactgccaa	gacatcttgg	ctttcacccc	164340
aactcaggat	ggtccagaat	ccagagctct	taagagagca	gatgctgaga	ggcacttaac	164400
ccagggctaa	gaccttccct	tggacggttt	cttggctttc	tactctgtcc	tctgtcccag	164460
tctgtcatcc	ccatctgtgc	ctaacagctc	tctgtggaaa	acatgaggcg	tatgagctct	164520
ctactttctc	cagcatccca	tgcccgcacc	ccagctcact	gtgtgccttc	atgttactca	164580
aatcttctgc	taggttttag	ggccccagg	tgaggctgtg	ggtttccctc	catctgtccc	164640
tcccttttac	caccaccact	aatcctcttc	ctcctcttcc	tctcttccct	cctcctcctc	164700
tttctnctcn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	164760
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	164820
aaattgaact	tacagaggaa	acacaaatga	gctgtaatga	cagaaaagga	atactatttc	164880
ccttgtagatc	agaaagctat	agattaaacc	ccagcccttt	ttgcctttgt	ttttcatatt	164940
tagtggttct	ggggactgaa	cccaagggaag	gacagctgct	tggaatgttg	gagacctctc	165000
aggcattgct	tatgctaagc	tgacaaaggc	tttctgaaatg	catatggtag	tatctattca	165060
tcttaagatg	cctataacca	tgggtggagc	aattctctaa	gagtctgtct	gaacttctta	165120
agcatgtgct	caaaaaacaca	cataaagatt	gtttactgaa	gcagtatttg	caacagtga	165180
acaaaaccac	acaacaacaa	aaagtactta	atagctcctc	agtaggaaaa	agaacaacaa	165240
attgggacac	atttatacta	ttctcgttct	atcacacgac	tattaagaag	attaagggtg	165300
gaggatgaaa	agcccctgca	gaacaacgca	cacaatgaaa	ttgaatagag	aataaagaaa	165360
aaaatatata	gtgtatttat	gtaaaacata	cataaacaca	taaaaagaat	agaaaactct	165420
ttgactgtat	atctttgaaa	agaaccagca	gagtgtctatg	gtttataaat	gaaaagttcc	165480
cccaaaattc	ctgagttgga	ggttcaatgc	agagttcata	ctcttaggaa	gtaactaaat	165540
caggaaatgct	ctgacttaat	caatggattg	accaactgat	agattaataa	tctgaaggca	165600
acacttccag	gaggcagaaa	caagggtgagg	cctaactgga	ggaagtttgt	caccagaggc	165660
atgtccttgg	agaggctacc	tcatccttga	tcttcccttc	tctccgcttc	ctggctgtaa	165720
aagtatttat	ttgactgatg	taactgtcat	cttggaggct	tactggctcc	atcagctaac	165780
ctaggccctag	ccctggaagc	ttctagcttc	catacaatct	aatccaagcc	tagaatgttc	165840
cagcctttag	gacttctgct	tgagatcacc	gtttctgtgt	ctttctgaa	tctagctggc	165900
tgattcagtc	cccctgttcc	gggctcaaac	tctctccccc	gatgatttta	ttcacaatct	165960
gtctttttct	ttggcctctg	aattgctctg	cttggctctca	aactaactct	agcaatcttt	166020
tctaactctct	tgtctccttc	acactctctt	gcttgttctg	tctttaactgt	gtctagtttg	166080
ttctctcttc	catccttctc	tgtaaaagctc	tcccggtaaa	cctgcctcct	cctccccctc	166140
tgtgcgcgtc	tactctccct	ctcagctcta	ctgcactgct	ctccaagctc	ctcctgtatc	166200
ctgtgctgca	ctctcttctc	cggtaaccac	tgtgtctccc	ttacgtagct	tccctttcct	166260
ctctcttctc	ctgagggttg	ggcagatcct	atcctgtcaa	acctttctct	gattcttcac	166320
tttgtctgcc	actcaattag	acatcacttt	caagcaggag	tgccctcctc	acaaaccaac	166380
tttaccttca	ttgtttcaaa	ttaaagggtga	gtactaaggg	tgtgtctctt	tttcagccag	166440
tgagagtaaa	gatgtgtgct	aataaggctg	agccaactct	agctagaaat	agtttctttt	166500
tctccataaa	taacagaatc	ttagggttca	caatacgatc	aaatatcctg	agacagctgg	166560
ctgccatgtg	gtgaagagaa	tgacagagag	taaggatgtg	ctcttgagg	ttcagatggg	166620
aataagcatt	ttcttgggag	ttggattaga	gtcattcctg	tgacactgtg	acaaaggact	166680
cgactacatt	ttgccatgcc	ttgagactgt	ggagggtgga	gcgatggatt	aatacactgg	166740
agtgaatttt	aaggcagcca	aggctgtggt	tggactgtta	ctagctacgt	ttagctggat	166800
ttatatgaag	aattgggaac	aaaaaagcag	agttagaaag	acagtttttg	cagaaggagg	166860
ttttctcata	tacaacttag	tttttatagt	tagcttttta	tgttgctatg	acaaaaatac	166920
cttatggaaa	ctgaagaata	aaaattttta	ctgaactctt	cttcacccca	gaacccgacc	166980
cctcccatct	agagattggt	cccggaaacac	tcttgaactc	ttcaccccag	aatgctttcc	167040
tgaactcctc	accctagagt	tccaacccctc	ccaactaaaa	actgttccaa	gaacattttt	167100
gagataaggg	cctcctaaaa	caacctcaaa	atgaaccggg	tacattgcca	aataatagga	167160
catgacccct	tagttacgta	gattcccttg	gcgaaccccc	ttgtcccttg	acagaacccc	167220
ctagtgtatg	aaacttgtac	tttccctgcc	cagctctccc	cccttgagtt	ttactatata	167280
agcctatgaa	aaatttggtc	ggctcgtgat	tctcctctac	accactagg	gcatgagttt	167340
cgaccccaga	gctctggtct	atgttccatg	tgctttcttg	ctgttgttct	attaaatctt	167400

gccttctaca	ttttgagtag	ggtctcagtg	tcttcttggg	tccgcggtg	tcccggggct	167460
tgagtgttg	agttaggggc	tcccttcggg	ggtctttcat	tttgggtgcat	tggccgggaa	167520
acagcgcgac	caccagagg	tcttagaccc	acttagaggt	aagggtcttt	gttctgtttt	167580
ggtctgatgt	ttgtgttctg	tttctaagtt	tgggtgcgac	gcagtttcgg	ttttgcggtat	167640
gctcagtgag	accgcgctcc	gagagggaa	gcggggtgga	taaggataga	cgtgtccagg	167700
tgtccaccgt	ccgttcaccc	tgggagacgt	cccaggaaaa	acaggggagg	accagggacg	167760
cctgggtggac	ccctttggag	gccaagagac	catttggggg	tgcgagatcg	tgggtttgag	167820
tcccacctcg	tgccagttg	cgagatcgtg	ggttcgagtc	ccacctcgcg	ttttgttgcg	167880
agaccgtggg	ttcaagtccc	acctcgcggt	tggtcacgag	atcgtgggtt	cgagtcccac	167940
ctcgtgcaga	gggtctcaat	cggccggcct	tagaaaggcc	atctgattct	ttgagttgct	168000
tgtgggtcgac	gcagagtcgc	cgccgtttct	ggtttctttt	ttgtcttagt	ctcgtgtccg	168060
ctctgtttgt	gtctactgtt	tttctagaaa	tgggacaatc	tgtgtccact	cccctttctc	168120
tgactctgga	gcattggaag	gaggtgcggg	tcagagccca	caaccagtcg	gtggaagtca	168180
gaaaggggtcc	gtggcagacc	ttttgcgctt	ccgagtggtc	aacgtttaga	gtaggctggc	168240
cacctgaggg	tgcttttgac	ttgtcactaa	tgcgtgccc	caggcgaatt	gtttttcagg	168300
aggaaggggg	tcacctgat	cagatcccct	acattgtgac	ctggcagaat	ctcgtccaat	168360
tcccacctcc	gtgggtcaag	ccttggaccc	caaactcttc	gaaactgacg	gtcgggtttg	168420
cccagctcta	tgacgcccga	aagtctggcc	catcagcacc	ccccaagatc	tatccagaga	168480
ttgacgacct	cctctggata	gactcccaac	ctccccctta	ccccctgccc	caacagccac	168540
ctgcagctgc	cccaccacag	ggaccaatag	cgagaggggc	tcaggggaccg	gcgggggaga	168600
ctcggagtcg	ccgaggccga	agccccgggg	aggaaggggg	gccagactca	acagttgcct	168660
tgccactcag	agcacatgtg	agagggccag	caccaggacc	taatgatctc	attcctttac	168720
agtaactggc	tttttctctc	tctgatttat	ataattgaaa	aactaaccac	cctcccttct	168780
cagagaaccc	ctctggactt	actgggctcc	ttgagtcact	tatgttctcc	catcaacca	168840
cttgggatga	ttgtcagcag	cttttgagg	ttctttttac	aacagaagaa	agaaaaagaa	168900
tcctcataga	ggcgagaaaa	aattgttctg	gagaggacgg	cacacccact	gccctcccta	168960
acctcgtgga	cgaggctttc	cccttgaacc	gcccccaactg	ggactacaac	accgcggaag	169020
gtaggggacg	cctccttgtc	tatcgccgga	ctctagtggc	aggctctcaga	ggagccgcta	169080
gacggcccac	caatttggct	aaggtaagag	aggtcttgca	ggggcagact	gaaccaccct	169140
cagtcttctc	tgagcgtcta	atggaggcat	ataggagata	cacccctttt	gaccccttgt	169200
cagaggggca	gagagccgct	gtagccatgg	ccttcattgg	tcagtccgct	cccgcattta	169260
agaaaaagct	gcaaaggctg	gaggggctcc	aagatcatat	gctccaagat	ttagtaaaag	169320
aagcagaaaa	agtctatcat	aagagggaaa	cagaagaaga	gaggcaggag	agagagaaga	169380
aagaaataga	ggagagggaa	aatagacggg	atcgccgtca	ggagagaaat	ctgagtaaaa	169440
ttttggccgc	agttgtgaat	gatagacagt	caggaaaagg	taaaataggg	ctcctgggca	169500
acagggcagt	gaaaccgcaa	ggtggcaaaa	agataccact	ggaaaaagac	caatgcccct	169560
attgcaaaaga	gaaaggacac	tgggctagag	attgccctaa	aaagcgggag	cgatccaagg	169620
tcctaaccct	agaagatgat	taggggaagtc	ggggctcaga	ccccctccct	gagcctaggg	169680
taactttgtc	cgtggagggg	actcccgta	gctctctgat	agacaccgga	gcagaacatt	169740
cagtactcac	taacccctta	ggcaagctag	gctccaaaaa	gacctgggta	attggagcca	169800
ctggtagtaa	attttaccct	tggacgacca	aacgagctct	tcagatagac	aaaaatatag	169860
tgaccacact	ctttctggtg	atacctgagt	gcccgtctcc	cctcttgggg	cgcgatctgc	169920
taaccaaact	aaaggctcaa	gtccaattta	cttcagaagg	cccacaagta	agctggggaa	169980
agggccctgt	tgccctgcctt	gtcctcaaca	cagaaaaaga	gtaccggttg	catgaagaac	170040
aaacccaaaa	tgcagctctc	tcagggttgg	taactgcgtt	ccccaatgtc	tgggcagaac	170100
aagcaggaat	gggggttggct	aaacaagtgc	ctccggttgt	ggtagaactt	aaagctgatg	170160
ccaccccat	ttcggtaaaa	caatacccca	tgagcaagga	agctagaaaa	ggcatccggc	170220
ctcatatcca	gaggttgctg	ggccaaggag	tttttagtgg	ctgtcagtc	ccctggaata	170280
caccacttct	gccggttcaa	aaaccaggga	ccaatgacta	tcgcccggta	caagacctcc	170340
gggaggttaa	caaaagggtc	ctggacattc	acccacagat	cccgaaccgg	tacaatttat	170400
taagctctct	cccacctgag	agaacatggg	atacagtcct	agacttaaaa	gatgccttct	170460
tttgccctgcg	tttgaccct	aagagtcagc	tccgtgttgc	ttttaaatgg	agggaccag	170520
agggcgga	gactggtcaa	ctaacttggg	ctaggctacc	acaggggttc	aaaaattccc	170580
ccacctgtt	tgacgaggcc	ctccatcggg	atcttgcgce	ttttcgcgct	cgaacccctc	170640
agcttaccct	actacagtat	gtagatgac	tcttggctgc	ggcgccctcg	aaggagctgt	170700
gtcaccaggg	aactgagagg	ctcctcacag	aactgagtga	cttgggggtat	cgagtttcgg	170760
ctaaaaaggg	acaaatctgt	caaactgagg	taaccttctc	ggggtatacc	ctccgagggg	170820
gcaaaagatg	gtcacagag	gcccggaaaa	agactgttat	gatgatccca	tcgccaacta	170880
ccccacggca	ggtacgtgag	tttttgggga	ctgtgtgctt	ttgtagactc	tggattccag	170940
gctttgcaac	cctagcagca	cctctatatc	ctttgactaa	ggaaggggtt	cctttcaagt	171000
ggaaagaaga	acacaaaga	gcttttgagg	ctatcaagtc	gtctctaatt	actgccccca	171060
cgctagcatt	accagacttg	actaagcctt	tcgtcctata	tgtggacgag	agagcgggtg	171120
tagccagggg	agtattgaca	caagcactgg	gacctgaaa	aagacctgta	gcctatttgt	171180
caaaaaaatt	agatcctgtt	gctagtggat	ggccacatg	tctgaaagct	attgcagcag	171240

tagccctgct	gatcaaagat	gctgacaaac	tgacaatggg	acagcagggtg	accgtttag	171300
cccctcatgc	cttagaaagt	atcgtgcgac	agccacctga	cagataagat	gacaaatgcc	171360
cgaatgacac	actatcagag	cctgctgcta	aatgagcgtg	taacctttgc	gccccctgcc	171420
atcctcaacc	cagctaccct	tctccctcta	acaaatgatt	ccgtcccagt	acatcaatgt	171480
atggacatcc	tcgctgaaga	aactgggacc	agaagtgacc	tgactgacca	accctggcct	171540
agagctccca	gttggtacac	ggacggcagc	agtttctga	tagaggggaa	gcaaaaggct	171600
ggagctgctg	tggtagacgg	gaaaaaggta	atttgggcaa	gcgctttgcc	tgaagggaaca	171660
tcggcacaaa	aggctgaact	tatagcgctt	atacaagccc	tccgagaggc	taaaggtaag	171720
atcgtaata	tctacactga	cagccgatat	gcttttgcta	ccgcacacat	ccatggggcc	171780
atctacaggc	agcgagggct	attgacctcg	gctggtaaa	acattaaaaa	caaagaaaaa	171840
attctggccc	tgttagaagc	catacatgca	cctaaaaagg	tagccatcat	ccactgcccc	171900
ggccacccaa	aaaggagaaa	acttgggtgc	caagggaac	cgaatggcag	acttagtgcc	171960
aaaaacaagt	gctcaagggg	ccatgatctt	aactgaaaaa	ggtgatccgc	ccaaaagccc	172020
tgaggatggg	aggtataaca	taaaagagct	atggtagacc	agtgatcccc	tcccatactt	172080
ttttgaaag	aaaaatagaa	ttaactcccg	aagaaggaa	aaaatttcta	aaaggactac	172140
accaattcac	ccacctggga	gttgaaaaaa	tgatgagact	aattaaaaat	tcccgatacc	172200
aagtcccca	cctgaagtca	gtggctcaaa	agattataga	ctcctgcaaa	ccatgtgcac	172260
tcactaatgc	aactaaagcc	tacagagAAC	ctggaaagag	acaacgggga	gaccatcctg	172320
gagtgtattg	ggaggtagac	tttactgaag	ttaaacttga	aatgtatggt	aacaagtatc	172380
tgtagtagt	tgtagacacc	ttttcaggat	gggttagagg	atttcccact	aaaacagaga	172440
ctgcccagat	tgtggccaag	aagatccctg	aagaaatcct	gccaagattt	gaaatcccta	172500
aggtaatcgg	gtccgacaat	ggaccagcct	ttgttgccca	ggtaagtcag	ggcttgccca	172560
ctcagttggg	catcgattgg	aaattacact	gtgcttaccg	ccctcaaaag	tcaggacagg	172620
tagagaagat	aaataggacc	ttaaaagaga	ccttgactaa	attagccatt	gagaccggca	172680
gaaaagactg	ggtggctctc	cttctctctg	cgctcaaaac	cccctggctg	tttcgggctc	172740
actccttttg	aagtctctga	tggaggacct	cccccttaa	tggaagctgg	tggaacatta	172800
gtttccgact	ctgacctgtg	cttacctctc	tctttgctta	ttcatttaaa	ggccctaaaa	172860
gtgattagga	cccagatttg	ggaccaactg	aaagcagcct	ataccaccag	gaccaccgca	172920
gtaccccaag	ggttccgagt	tggagacaaa	gtcttgggtc	gacggcatcg	aaccggtagc	172980
cttgagccac	ggtggaagg	accctatttg	gtgttactga	caacccttac	tgcggtaaaa	173040
gttgacggaa	tcgcctcctg	gatccacgcc	tcccacgtca	agagggccgc	cagtcaagat	173100
gaagaaaacc	acgacgacaa	ttggacagtg	gcagtcactg	acaatcctct	taagcttcgt	173160
ctgcgcgcga	ggcgccactc	tagacctagg	gaaccttaac	cctcatgctc	caattcaaca	173220
gtcctgggag	gtgcttaatg	aaaaggaaaa	acttgtatgg	gcaaccactg	cagtcacatc	173280
cctctggatt	tggtggcctg	atctcacgcc	tgacatctgt	aagttagcgg	caggatcccc	173340
caattgggac	ctctcagatc	atactgatct	tagcaaccac	ccccctgagg	agcgggtgtg	173400
cccaaatggg	atagggagca	catatgggtg	ttcggggcag	ttctaccgag	ctaattcttag	173460
agctgcacat	ttttatgttt	gccctgggtc	gggtcagagc	aaaaggcttc	aacaaaaaatg	173520
cgggggggca	tcagattact	tttgtgttaa	atggacatgt	gaaacgacag	gagatgctta	173580
ctggaagccc	tcctctaaat	gggacctaat	cacggtaaaa	cgaggtagt	gctatgataa	173640
gtcaaacgaa	ggagaaagaa	accctataaa	atatcaagag	agtgggtgcg	cttttaaaaa	173700
cagagcacc	tcaggaccat	gcaaagataa	atactgtaac	cccctacgta	taaggttcac	173760
cgagaacgga	aaacaacacc	gtctaagttg	gcttaaagg	aataggtggg	gttggcgagt	173820
atacattcca	ctaagagatc	ctgggttcac	tttcacgatc	agattgacag	tgagagacc	173880
ggcagtgcac	ctcgtagggc	ccaacaaggt	ccttaaaaa	caggggcccc	ccagtcgtac	173940
tggtcccccc	aaaggtcccg	actgtaccag	ctccaccaac	tccacagccc	aacacagtgg	174000
taccctccct	aggaactaat	actctcctca	taaagcctac	cttggcttcc	ccaccgcccc	174060
taggaacaga	ggaccgtctg	gtcagtcctg	tccaaggagc	tttttttagt	ctaaatagaa	174120
ctaaccctaa	tatgactcaa	tcatgctggt	tatgctatgc	ctctagcccc	ccttattata	174180
aaggaatagc	tcagatcagg	acttataata	ctacttcaga	tcattctcaa	tgcccttggg	174240
gaaaaaacag	aaagttgact	ctagcagcag	tttcagggaag	agggctttgt	ctgggcccgg	174300
tacctcagga	taaagggcac	ctctgtaatc	agaccagaa	catccagtct	agcaaaagcg	174360
gtcagtatct	ggtgcctccc	ctagacacag	tgtgggcttg	caataccggt	ctcactcctt	174420
gtgtgtctat	gtctgttttt	aatagttcca	aagattttctg	catttttggt	cagcttattc	174480
ccagactcct	gtatcatgat	aatagttcct	ttttagataa	atttgaacat	cgggtccgct	174540
gaaaaagaga	accgtttacc	ttaacttttg	cagttctatt	aggattggga	gtagcagctg	174600
gagtaggtac	aggaaccgct	gccttaatta	agaccccccc	aatactatga	agaactacgt	174660
gcagttatgg	atattgatct	tagaactata	gaacagtcta	taaccaaaatt	agaagaatct	174720
ttaacttccc	tgtccgaagt	ggtgctgcaa	aatagaagg	aattagactt	attattcctt	174780
aaaaaaaag	gactctgtgc	tgccctaaaa	gaagaatgtt	gtttttatgt	tgaccattca	174840
ggagtaatac	aagattctat	ggctaaactt	agagaacgcc	tagatatacg	taaaagagaa	174900
agaaaaagcc	aacaaagatg	gtttgaaagc	tggtttaata	agtccctctg	gtccaccact	174960
ctcctctcca	ctatagcagg	acctttaatt	acacttatgc	ttttgcttac	ttttgggccc	175020
tgcatcctta	ataagttagt	agctttttatt	agaaaaagga	taaacgcagt	ccaggttatg	175080

gtactaaggc	aacaatatcg	ggtccttcag	gaggttgaaa	actcgctcta	agattagagc	175140
tatctcctaa	aagaagtggg	gaatgaagaa	taaaaatfff	tactgaactc	ttcttcaccc	175200
cagaacccga	cccctcccat	ctagagattg	ttcccgaac	actcctgaac	tcttcacccc	175260
agaatgcatt	cctgaactcc	tcaccctaga	gttcgaaccc	tcccaactaa	aaactgttcc	175320
tagaacatft	ttgagataag	ggcctcctaa	aacaaccgca	aaatgaaccg	ggtacattgc	175380
caaataatag	gacatgaccc	cttagttacg	tagattccct	tggcagaacc	ccttgtcccc	175440
tgacagaacc	ccctagtgat	gtaaacttgt	actttccctg	cccagctctc	cccccttgag	175500
ttttactata	taagcctgta	aaaaatttgg	ctggctcgtc	attctcctct	acaccactag	175560
gtgcatgagt	ttcgacccca	gagctctggf	ctatgttcca	tgtgctttqt	tgctgttggt	175620
ctattaaatc	ttgcctttcta	catftttgagt	acggtctcag	tgtcttcttg	ggtccgcggc	175680
tgtcccgggg	cttgagtgc	tgagtgaggg	tctcccttcg	ggggtctttc	aaaactactt	175740
cagaggaaaa	atgtattctg	cctcatgggt	tcaggggggt	tccctcagca	aattcaggga	175800
agacaagatg	gaacagctca	acctgctggc	aggagggtgt	gggaaaggac	aagtgttcat	175860
tgtgtgggtg	acaggaaaca	gagagctgcc	tacagtctta	caggcctacc	accactgacc	175920
tacctctgtc	cgtcaggccc	tacatcttaa	aggatctaca	gtttattaaa	agaacactac	175980
cagataggaa	ccaagtatca	aaccaccagt	ttgtagggga	taaaaatata	aggaacacat	176040
ctcaatagga	gtgtgttcca	ggatgtggac	aaggagaaca	cagttgttta	aaagcttaac	176100
gctggccagg	agagctgcac	acctttaatt	ccatcactcg	taagaggga	gcaggttcat	176160
ctctgtgagt	tcaaggcaag	cctgggctat	acaattctag	attagccaga	gctacatcgt	176220
aggagcctgt	ttcaaaacaa	acaaaaccaa	accataaaaa	agcatttctg	aggctttggg	176280
tttaatcccc	atgacctcaa	atagccaaac	agctctcctc	agtccaaacc	aaactgcaaa	176340
attggagcta	gtgagatggc	tcaacatatg	aaagtccctc	ccaaaaatat	tgacaactgt	176400
agcttatctc	tggggacaca	cataatggga	gaggaccaat	ttctacaagt	taccctctga	176460
cctccacaca	tatgcctccc	acaaataaga	aaatatatat	aataaaaaaga	aagaagtcta	176520
cagctgcaca	tggtcatgca	tgcctataat	ccagcactcc	agaggctgag	gcaggaggat	176580
tattagtttg	agatcgcata	gcaagcagta	ggctagacag	ggctacatag	tgtaaacctg	176640
ccttaaaaaca	caaaaatcaa	ttaagcaaca	ataacagtaa	caaccacaac	aaaaacccaa	176700
aagagtactt	tgtagtaagg	acaataccaa	aaatgttccf	ttaaggacag	ttctggaaac	176760
agcaatagcc	ttccgagtgc	tcagggatgt	ataaataact	agaaaaactc	ccctggagaa	176820
atgagcacca	gggtacactg	ctctcagagc	tgcccagaaa	gttggtttatc	ctggattcat	176880
ttcagccttc	ctaactgctc	aggcattcag	aggtcacttc	tgtagtagcc	aatgtctaaa	176940
aaggctaaac	tactgctcag	catggctgtg	gtacttgcca	ttatcatttt	gtgactgggt	177000
ttgtagttaa	gcagaattca	agagttatag	catcatgaaa	gtttccacca	agttcctgat	177060
ccagtcaccc	cttaaaagggt	ggatgcacca	agtgcctttg	gggtgataaa	ttatatcaaa	177120
ataatgggtat	tccaccctaa	tcccaaaaga	ctctggcca	tctcataatg	taaaaatgctg	177180
agccatcgca	ccagcccatg	gccttgaact	cttgatggte	ctgtctcagc	ctgtgtttgg	177240
attataaaatc	tgttgggtgag	gtattccttt	gctgataata	caagcaaatt	cttcaagctt	177300
ccatcctaga	ctgaagacca	gcagctctcc	aggagtccct	aatgcagact	ggcccagctg	177360
ggacatttag	cctcatggac	tcagccgcta	ctagattcgc	aacctattca	gacaagccac	177420
tgttgagcta	ccagacaaat	actatgtaag	ccaatcccat	tttaatacac	atattcatct	177480
gggtgtgtgg	cacacacctc	tactcccagc	acgcaagagg	cagaggcagg	cagatctctg	177540
atttcgaggg	ctgggtctata	gagtgaattc	caggccagcc	agggtacac	agagaaaacc	177600
tgtttcaaca	aaacccaaaac	cgtaaattca	ttctatcagc	tctatttccf	tagagaattc	177660
taatacatgt	gggtaccagg	ggttgaactc	aaagtcttca	tgtttacgta	gcaagtttcc	177720
ttctgctagc	ctagtgaagc	tgaggcagg	acagccgggt	ctttactgct	ccttgcaaat	177780
ggctccctcg	agcttttccf	tgagagccta	caaagaaactc	tttttttctt	taggtctcca	177840
ggttttggte	ttaagagggt	ctggacttgg	atctgtagct	gtcatatcac	agacattcaa	177900
catctggcaa	atgtcttgac	aaaggagatc	acttgtgttt	gctgcagtg	ccctctctggc	177960
tgtgagattt	tgctccctcac	cactgcagga	ctgcagatct	attctgcctt	tttagttgac	178020
ttttcatttc	tgagaactgg	ggaaaactga	ctttgtattt	gggctttgaa	tttgtccatt	178080
tgtcaatcca	tcacaccaga	cctaaccaac	tgccaagagt	tctgctgact	tttgttttct	178140
ctagggtggf	cactttgctg	ggctcatcct	catccttggc	ctgcagttta	tccccaggaa	178200
agaaaatggc	taacgactgc	taagaagcag	tctttccctc	cagaaatfff	agtcctatcta	178260
gaccttgctg	cagtctgaag	tctttaaaat	gtgtttgtta	tggtagaata	ttttgagttg	178320
ccttaggagt	attgcttgct	gtcacctatc	atattctatc	aggaagcaga	cgtcccatft	178380
accaaattgtg	aagaaatatg	gcatacaatc	ccactgcaaa	aagtgtaaat	aaataataaa	178440
aaaaatagatt	tattacagag	tgcaaggga	aagaaaaaaa	tcagccagtt	tcagaattgt	178500
aactggacaa	atgttggtac	agttcatgaa	gaggttctac	aaaatggctg	ggggtgggaa	178560
cataatgagt	tagtttgctt	ttttttttct	ctttcccttc	ctttcccttc	cttacaagg	178620
ctcatgtagt	ctatggctc	aaactcacca	ctgtaaatca	ccttgaactt	ctgatccctc	178680
tgcacgttg	caatgtaagc	atgtgccacc	aggctggct	cacacatttg	gtttttcaat	178740
acagaatagc	tctgtgatga	ttaacttcaa	tcatcaactt	gacataacca	agaatcgtct	178800
gaggaagagt	ctcagtgact	gggtgggcta	agggcatgct	cataagggat	tatcctgatt	178860
gttaattgac	atggaaagat	caagtccatt	gtgagcagca	acacgcctg	aacagaagtc	178920

ttctgaagta	taagaggaga	aagcttgatg	agagcaagca	ggcaagcaag	ccaggatcca	178980
cgtgtttatt	ctctgtctgt	tcttgaccgt	agatgtgatg	gctgtcttgg	cttcctggga	179040
aacatgaact	gcaccctgga	attgcaaggc	aaacaaacct	tttcctcttc	caagttgctt	179100
tatgctaaga	tattttatcg	cagcaataga	aatgaaacct	agaacaggcc	cataactgcc	179160
agctttggaa	ctgaacctaa	ggctgttata	attcactagg	atagggacca	ctggaagtga	179220
atctgatttt	gatggtaaaa	tcatgtgttt	gtttctggat	atgatagatt	tatcaatttg	179280
agactcagaa	aagaagttag	gacttgaatt	ccgtttttaga	gacattccag	agaaaaactga	179340
tgtcattgtt	ctgaatgtaa	gtgcctcagc	tgaaaataca	aagagtacag	ggaagaaagc	179400
ccaggctaga	atctgaagga	actcctctat	tttttgtttg	cttgtttgtt	tgggttggtt	179460
tttgagacag	ggtttctctg	tgtagccctg	actgtcctgg	aactcacttt	gtagaccagg	179520
ctggcctcga	actaagaaat	ctgcctgcct	ctgcttccca	agtgtctggga	ttaaaggcgt	179580
gtgccaccac	accaggctag	gaactcgtct	attacacatt	aacacccttc	tttaattaac	179640
tgttctgtcc	aatgtaccaa	atagtcaatt	gattcctgtt	tatttaccac	atgtttctgt	179700
tagtaacca	gaataactta	tctagccaaa	gtctgcctat	tagccatatt	ttcatcagtt	179760
cccaaccatt	tttggaattc	tgtgagggga	atccacagat	gctgtagacc	gctttagaca	179820
tttttcagct	tttttcaagt	tgcaggctcat	gattcagtg	gtcatgaaat	taatttagtg	179880
ggttctgatt	agcatttcaa	aatgaggcaa	gcagagggca	tattgtcaca	gcacagcaca	179940
tgcggtaaag	agccacacac	tcttgcttgg	aggcttagtc	agtttctggc	tctaaacgcc	180000
ccaggtttgt	ttctctatcc	taggcctctc	tcttaaattc	caaacatagt	tagacattac	180060
cattggggca	cgtgcaactc	aaacacggag	tgtgactcct	ttccccatct	gcggttccca	180120
gatttggcaa	tgtcacccctc	ctccctctctc	cctagggtca	gttttacctc	tcacactcca	180180
caacacaaca	cctctcatct	caagaattgc	cctaggggct	ggtgagatgg	ctcagaggtt	180240
aagagcaccg	actgctcttc	tgaaggttct	gagttcfaat	cccagcaacc	acatggtggc	180300
tcacaacccat	ctgtaattgg	atctgattac	ctcttctggt	gtgtctgaag	acagctacag	180360
tgtactcaca	tatattaaat	aaataaatct	aaaaaaaaaa	aaaaaaagaa	ttgccattaa	180420
atgtacctca	gagtccaaat	gcttcttctc	cccctgacta	cactcacgct	ggcctgagtc	180480
cattttctta	ttgaggttac	tgttctctctg	cttctaccct	ggctccttct	gctgcctatc	180540
cttgacacag	cagacaagca	gttcttttaa	gcagggtctca	ggaccagtga	gactgatcgg	180600
ctctgtggc	acttctgtcc	atgactgatg	atctaaagg	aagcctagaa	cccacgaggt	180660
agaagcaaag	gacctactct	ccaaagccgt	cctctgacca	ccatgtgtaa	actgcacatg	180720
tacatgcatg	cacatggtac	acacacatac	acagaagtaa	aaagagattt	aaattgaaaa	180780
tcattaaaaa	gaaaaatcag	ggctcagcaa	actttccgtg	tagaaaaacta	gagtacttag	180840
gctttgaaag	ccaagaagtg	gatattaaat	atagttattc	attatagcag	agatttctaa	180900
aaccttttga	caaaaactaa	aaatataaca	gagtgatttt	tttttgaat	gtaagtttac	180960
taatggcagc	agtgggatta	gtttcttttt	tagattattg	ttattatttt	tattaattat	181020
tagtggtttt	gtgtttattc	atattccaca	gcatgtgtgt	ggaattggat	ttctgcttcc	181080
acctttgtgt	gggtcctaga	gattgaaact	aagtcatcaa	gcttgacag	taggttggtca	181140
ggcttacaca	gtaggtggtc	aggcttgat	ctttggaagg	caagcatttt	acttctctgt	181200
ccagctcact	ggccttcttt	gtttaaaaaa	agtaaaaaaa	agtccttttt	tgtttattta	181260
gggttcatgg	ccagtgtctt	ttatctttaa	atcaactgca	aactttttatc	tggtaaaaag	181320
ccatccttag	ctgtggtcct	aggagaaaaa	catacagttg	gatggcttta	tcctgcaggc	181380
ttagtttgat	catctctctt	tgaagatata	atcagctcac	atcacactca	agcctctgcc	181440
aacgagtttt	ctacttctgt	tcaacaaact	acccaagctg	agcagctcca	aacaacagcc	181500
agttatgatc	ctcacagtcc	ggtgggtcag	aagcctaaag	gggcgtggct	acctcgctgc	181560
tattgcctga	ccctgctcgg	tgcattccaca	ttcacatcct	ttcctggtga	gtgtggttct	181620
ttgactgggt	ttgttccaat	ttttagtata	tgtgctgctg	aaacaatctt	tttgctctctg	181680
cctccagact	gcagggatta	atgttcttga	ctgccacaga	gcactaatat	ttactgaaca	181740
tgtgatcatg	tgggtctcag	cactcttgca	cccaaggctc	ggggaacatg	gaggaagagg	181800
gggtggaaag	attccaagaa	ccagaggaag	aagaaagtca	gaggtgagac	tgcatctcct	181860
agaaatgtca	gggacatttc	tagacctctg	aagttctcaag	aacaaggcct	gaaagtctta	181920
tttataatag	ttaacctgaa	aggggaaaaa	attcttacag	gggtccaacg	ttagacaaag	181980
aactctaagc	aactaaggaa	tgttgggggg	ggggtagtct	tccccaggga	acactcctct	182040
acccttcaag	ccccacccaa	gctggttata	caaaacaaac	tggtcagtcc	tgaagccata	182100
tacgcacaag	taacatcata	tggatgggca	gattgcattt	aggaatacac	acatacacac	182160
acacaactta	aaaagagagg	ccatgaattt	aagagagagc	aaagcaaagt	gggaaggggg	182220
acatgggaag	gttggaggca	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	182280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	182340
nagagagaga	gcctattatg	tctgttggtg	cttctaatca	ttagaaaacc	actctcttag	182400
gctgagtcag	aactagccta	gctgagacac	tgtccaccca	ctgtcccgga	gcaaggccat	182460
cgtgtcccca	gatttgcctt	tgggggcctt	tgaatgaaa	gtcaccagca	ggctccggaa	182520
gctgctcat	gctaactcag	tgaagttgtg	aaaatagccc	tgcagtgtgt	cctgggctctg	182580
cagctggggc	agagccacta	aggggagctt	ggtccttttg	agcagagtta	acagtcatca	182640
gtgctttttt	ttttttttta	aatgttccct	gcttttaggt	cagtgtgtgtg	cgctacttct	182700
aatccttgca	ataggctgca	agacaggcaa	gaatatcatc	cctgttttgc	cctcaggcaa	182760

attctgaagt	ctggcaaatg	aaagatgtgg	gatttgaaca	cagacttgtt	tggccaaaga	182820
attctcactc	tacttctgcc	tgtgccacct	tcctctcatg	cacggggagg	ggaggggagc	182880
ccacctccca	tgctcagggg	ctaggaagtg	gggagaagat	ggatgtcctc	aaagcagggg	182940
gagaatgaag	tagaagccag	cttcaaattc	aaactaagca	atgttttatt	tccattttccc	183000
tgaacataaaa	gttcagttac	atttggttta	aaaaaaaaaa	tccttacaca	actggttctt	183060
gagaaatgtc	aagtgtctaca	attcagtgga	tgtggatgaa	acaatcaaaa	tggtgaacac	183120
ccccaacag	atacaaacct	tcatcaaagt	ctcttccaaa	ggctgggtct	gaaaagagcg	183180
actcatgttc	cagcccagtt	ggctccttct	catgtgagct	ccgacttcca	aagactgctt	183240
gcaccaggag	gaaatataat	agatgtcctt	tttaaggggg	gtggggtctg	tctgacaacc	183300
tcccacagtg	actgtggata	cagcccagtt	agtagagttc	tcgcctagca	agcgtgcggc	183360
cctgggcttg	agatctagca	ccctaaggca	tagtggtaga	tgcccatgac	cacagcactt	183420
gggacgtaga	ggcagaagga	tcagttcaag	gtcagatgag	gggtggggag	gcatttcttt	183480
aatgcccagca	tttgagaggc	agaaacagat	gaatttgtga	gttcaaggcc	agcctgggtct	183540
acagactgag	ttccaagaca	gccaaaggcta	cacagagaaa	ccctgtcttg	tcaggaaaaa	183600
agatagtggg	agagaattca	aggttatctt	ggactgcata	agactttgat	tccaaaataa	183660
acaaaaatgg	agcatgaatg	cttgcaactg	tggacaatat	tgggttcata	catattctgt	183720
tttgtcacct	acataccaat	tatacaaatc	agattcagct	gggcccactg	gtgcatgttt	183780
gttcccagca	tctgggaggt	agagatgggc	agagctctat	gactttaagg	ctagcctggg	183840
ctacaaagta	agttctagga	cagccaaccc	tacagagaga	tacactactt	ctaaatcaat	183900
caatcaatcc	atcagtcaat	catgggctgg	agagatagct	cagtgtacaa	tagtgccatt	183960
cttccagagg	acctgggttt	gattcccagc	acccacatgg	cagctcacag	atgtctgtaa	184020
ctccaatccc	aagggacatg	acaacttcta	ctggtctctt	tggtcaacag	gcatgcacgc	184080
agcatacaat	atatatatgg	gtaaaatgct	atatatatata	aaatcagatt	cacaaatcaa	184140
gtacagaaaag	agatcaacta	taatgaaaca	accataaacac	atactgtttt	aaaagctacc	184200
tttctgtct	gacatctgac	ttcttgcgtg	accagtgcc	tccacaatga	gacaagcaga	184260
tggtgtagtc	cctgtgaccc	aggattaggc	aaccactgcc	cctgggaagc	cattcctagt	184320
aacaggatca	aatcccacag	tgctccactg	taccccgcga	gaggacgtga	agcctccctt	184380
ccccggctgt	ctgtgcagca	cgcgccagat	gcttgggtgc	actgagtgc	catcggagcc	184440
caactgagga	gtgcctcagt	gctcctcagg	gcagtgtgca	attgaaactt	gcatgtgatt	184500
tctggaattt	gtcatttcat	atttccagac	tccagctgac	cttggataac	agcacagagg	184560
gcagaactgc	agaggaaaag	gggcaactact	gtgactgtta	tcccctgcat	gattatagaa	184620
gggtctgtgc	tttctcttac	aaatcattgc	ggcctctctt	cacttccctc	ctttgaagtc	184680
gaaaaaaata	gatgcatcct	cacagtacag	gatggcagggt	acgaggcggg	ttccgggact	184740
gaggcaggac	tcataatgca	gcctctttcc	ctcaactaca	ccgccgcgcc	tgcaagtttc	184800
cctctgatca	aatcagtttc	aggcctggaa	agaaacggcc	actcaggctg	gggatgtggt	184860
tccattagag	gcctgcatgc	acgaagctct	ggatttgatc	ttcagcacgg	gcataagccc	184920
agtatggtag	tatctgttta	gcatggtgtg	gaggtatacc	tatctctgca	cagggagacc	184980
agaagttcag	agttatcctt	gcacttacag	taagttcaaa	gctagcttgg	gcaacatgaa	185040
gttttgtctt	aacaaaacga	aacaagaggg	gcgggggaga	tgcttcgctg	tgctgagtca	185100
tttgctgcca	agtttgatga	cctgagtttg	gtcccttgag	cttatgggtg	aaggagagaa	185160
atgacgcttg	aactccacgc	acatgccaca	gccacgcac	gaatgtgtat	acacacacac	185220
actaagtggg	taaatgttaa	aaataaataa	atcaaaaggaa	tggccactca	aaatctacca	185280
tcgttgggaa	gggaggggaa	aaggcaggcg	agggagatag	ataaccctga	tatgaacacg	185340
gaaagagcca	gtgtgccacc	aaagctgccc	agtgtgccac	caaagctgcc	cagtgtgcca	185400
ccaaagctgc	ccagtgtgcc	ccagacttg	attacagatt	tgccagggga	185460	
cacaggaggc	cagcaggagc	agccagggtt	cacctcagag	gtggagccac	aaacctggaa	185520
atgaaacgtc	tttccctttc	ttcagaccac	agcagtgcac	gctgtcctgc	agagtctgga	185580
gggtggcag	ggctcatcca	ctctagtgtg	cctgtggcca	gaacaggcct	cagtcaagg	185640
tgcttttcca	aggtcttagt	gtctaattaa	ggttagcagc	caaattggag	agagaagggt	185700
gctggacttt	actctgctgt	aaggactttg	ggcattgttc	cattccgtga	tcaaatacca	185760
ctggctctgc	caaccaccat	gtcagtgggt	cttcagaggt	agaagaactc	atcctttttt	185820
gagaggtttg	gtctggtcct	tgtctaatac	aaaatgcctg	gggcaccagg	ttaatgtcaa	185880
ctcaaaggca	agtgtgtgtc	cagcatgtgt	ggaatcctaa	gttcaatacc	catcagagcc	185940
ccaagcccta	gaggacagac	atgctttaa	aaaaagtcac	gctttaaaaa	aattctgttg	186000
agggggctgg	agaaatagct	cagcagttaa	gagcactagc	tgctctttca	gaggaccag	186060
attccgttcc	taacattttc	atgggtggctc	aaagctgtct	ataattcaag	tcctagaggg	186120
gaatctgttg	ccctctctgg	ttttctcagg	caccaagaac	acatgtgtgtg	caaacatata	186180
cgcaggcaaa	acactcatac	atatgaaaca	ttttttataa	acctgctcgg	atgtgggtggc	186240
tcatgccagc	gatgctctca	gcactcagat	ggcagaggca	gggggatttt	gtgttgagcc	186300
cagcctgagc	tatagaatga	gatgctgtct	caaaaagaaa	aaaaaaacaa	aaaaaaacaa	186360
aaaacaaaaa	caaatctgtg	ggcttaatac	ttcctagcca	gaaggagctg	gctccagcaa	186420
caatggatcc	ctagagctct	gctttgcccc	ctggtctgga	gagcttgctc	tagaaaggaa	186480
ttctccatag	accacatttc	tattttgggg	accactctgg	gcctgcacat	ggaaaaatgg	186540
agagttgagg	tttacatggt	catttttttt	gttggttacag	taatgagagc	tttgaaatga	186600

tcacaaaagg	aaaataggaa	aatatgcctc	ctaaaaagag	ccgaggcaaa	ttattatagc	186660
aaccgaatcc	taaaaggaat	gttcaagtaa	aaaaaaaaaa	atatggctca	tgccaagtgtg	186720
ttatggcaac	tgacatttaa	aagtaacagg	atttggggcaa	gtttctttct	ccacctctc	186780
tgcaaaagtct	tgcaaggaact	tcatgtctaaa	attatgtcttt	aatttttaatt	agccaaatag	186840
gtcataaata	tagctttattc	ccaaatcctt	agaattttcta	ccctgcaagg	agctgaacta	186900
ctaattgagga	aatattttcca	caaaaaccca	cttaccattta	agaaccccca	tccgatattt	186960
ttctaataata	atztatcaat	ttaaaccacat	tgcataatgt	gccactctgt	agcattccat	187020
taaaatgata	aatagcaata	gtggtgaggg	gtgggggggca	aaaaccagga	gattaaacat	187080
atccaaagca	gtgtagctat	atttaaatac	ctcaatccat	tgtagaggga	aacacactgt	187140
tctcatccgc	agatacagtc	tagactcaga	gcagcatatc	cttgactgta	agggtatttaa	187200
taggacagac	gaagggggggc	aataagaaat	gacaggaaac	ttcagaagaa	ataaaatttc	187260
tatttagctt	tgttataaga	ttacatcaaa	gcagttcata	tgttttaatc	tggggaggaa	187320
aaaaagcaac	tacttggggt	ttgcgcctgg	gggctgcctc	tgtgtactga	accagacagt	187380
ttgcataatg	aacaatttttc	attcaatcag	gatctcagca	gagatagctc	ctactcaaag	187440
gaacccggca	caggctcata	gtttttatct	cccagctcca	cctgctggag	aaaccttgta	187500
ttgcagggag	agaaagcagt	cgggaggcat	tgctctagt	gctgtgtacc	taaagttaca	187560
gcttgactt	taaacagttt	ctctctggag	gttgaaagg	gctctgtaag	ataccagagt	187620
ggattgtct	caaagactct	cggactcctg	ttacaggcaa	gtaaggctct	agcagatggg	187680
agcatggatc	tccggccctt	ctcactgctt	tcttgaatca	gggatttaga	aattgctatt	187740
tgcataccag	gaggactgaa	gtttggctcc	cgggtgaccag	aggacaagg	cattgttttaa	187800
aaccacccaa	actcatttcc	gacttggttg	gtcaaatttt	caagtttccc	agcagtctaa	187860
ggattcataa	aataaggcag	aggcagagaa	acgggaggtg	tgtgtgtgtg	tgtgtgtgtg	187920
tacccaaaat	ggaaactgcat	tttcatgcac	aatagaaaaac	ttaaagactg	aaccaatcat	187980
tttggaanaac	tgccacagct	gacattggct	agaggaagga	acggccagg	cgagccagct	188040
gcaccaagac	ccagggtgta	ggcctaatac	gcctttatcc	gaggggttag	tgaggctccc	188100
gccgctcacc	aatcccggct	ggagccgcag	aagagctctc	ttcacttggc	tcagtcccag	188160
cacagtgcga	ctatgctctc	ccgtggggag	gccgctccgg	gagggggagc	gacatcaagc	188220
tttgtgaaac	tgttttcgaa	aacctgggat	gatcatttaa	atgtttaaaa	tatgcacatg	188280
gtaattcaaa	actaattacc	ctgagcacat	ttgaaacatt	tatgccatca	tcttggatcc	188340
tgccactga	ttgtgcgctg	cagctcactc	tggtgtttct	ataaactgct	tcagcgattt	188400
taacttccag	gctaaatcag	gcagccacag	gcgctgcctc	cagccctggg	ttgggtggaga	188460
gacccccatc	cctgacttcc	aggcgaggag	gcggcccggt	tctccagaga	gccgtttgtc	188520
agggctctgt	agttctggct	gccgaattat	tgctcttatc	cgtgttcata	attctcatct	188580
gcattatttta	atttaggcta	gaatgacctc	tttccctccc	gagtcttctc	ccctcattcc	188640
catttctctc	tcttcaattc	gtggccccc	ttttctgatt	ggtccaaata	tatagacaaa	188700
tatccttgat	cgtcccaccc	cacttggtta	catcttcatc	tgggagccaa	tgtggtgagt	188760
tttctgggtt	tgcaagggtg	tcagggtccac	cagtcacctc	aagggtgtgtg	agagaggtag	188820
accaacatga	gcggcgcaca	gccgccatca	ctgagagagc	acgtgccctg	cagctcaggg	188880
acaggcgtg	acacaccggc	agacatgtgc	acatgcgctt	tccccagcaa	acctgctgtg	188940
cagagtaatt	aggcctaggc	agttcctgaa	gcaaattcat	ttcccccttt	tccagaataa	189000
aatgagttct	cttcccttgg	gggtgctaaa	ccagcatgcc	agtggctaga	agcctgagat	189060
gggtgatgtg	gctgaaacca	tttctgcagc	caagcctgtg	ggcagaagct	aaccttgggg	189120
tggggagctg	cagtcgggaag	aggcacaatt	ctgggatcaa	gaaatgagca	ctgggtttata	189180
ggtacatccc	cagaaataga	cagatgaggg	ctgcctcctt	attagcgctt	tgaagatgcc	189240
catggcgggt	ttttagacat	ttaggaatat	aaaagtaggt	tggattccca	cagtcagctg	189300
aagtttgaca	gagtgatatt	accgggttta	actagagcca	ttaagagact	cttcattatc	189360
ccacaccacc	gccacccaag	ttatcacatg	agccataatg	caagagaatt	ttcattccat	189420
caacaagaga	gggagccggg	ctatctttgt	ccaaaggaaa	tgagcagccc	agcgtgaagc	189480
ttgtgaggaa	ttgagtgtac	aacactccaa	taacatcccc	tgaggatttg	cctctgcgat	189540
ttagtcgggtg	aagcaggggt	aactgcgctc	gagcagctcg	cctgtgtacc	tggcttgcaa	189600
gaacaccagc	tcgagggaaca	ccaaaaaggc	cgattaatga	caaaggacac	tcatagaggc	189660
ccgaattcca	cagggcttaa	gtattaagcc	ccaaagaaat	caagggtctag	gccattctcc	189720
tggcgctcag	caatctcatt	tattatttct	ctacaaagat	ccaacactca	atttcccagg	189780
tatcccctgt	atctgactca	cattctcctg	ctcagtaagc	catcctgggt	tgaacggggc	189840
ctcccctcct	cctgcctatg	catgctttgc	gtcttccaaa	cgacagctgg	taatttgcga	189900
gacccccctc	actggactct	ctcaccaccc	atacttggaa	ctactccttg	gaactacttg	189960
tttatcaagt	gttctgttgg	tgagccttct	cttgcattaa	agctgtgaga	aggaaaccaca	190020
gtttctatatt	cctttacatt	tcttgtagcg	tctcacatgg	gagacaccca	ggttagatat	190080
actgagggtc	ctggtagttt	tagagttgga	gttagatgac	ccagcaacat	gccttcccc	190140
accacgcacc	aagcaaaaat	tgcaccaccc	atgttctcag	atgttctcgg	catcttataa	190200
ctcgcccaaa	ccagatttta	ttgctcctgc	tgtaaagtgt	atcttctcta	agcctcactt	190260
aaaagctacc	acttggcaga	agatcaagtc	agaagtgcag	gctagcaggt	gacgggtgagg	190320
acagggcggg	atggggcggg	taggggtggag	cgaataattg	aagctccaag	agttaccagc	190380
tcaatatttta	acctaactgg	taatttgctg	tgacaattac	gccatgaagg	gaacgctgcg	190440

actatgcaag	aatgttgctc	tctaattaag	agggtctctg	atttcctagt	cacccgcact	190500
ttaataaacac	acagaatgag	ccttggtctc	gggagctaaa	ggttcatta	ggagcacggg	190560
cagcatatgg	ctgtgcacat	aggccgtgag	tgatgcagcc	cagttaagcc	cgctaacc	190620
ttcaattcgt	cctcagatag	agcccagaga	gcgcggctca	ggccctcacg	ccacgagccc	190680
catttgactg	acaggcatct	tcccggaaag	cctgcgcgtg	cctacactgc	aaatggacct	190740
gcttcccaca	gcccggcttt	caaccaggaa	ggcttggcgt	gggtctgac	cttcaagagt	190800
aactttaata	aggattttct	cacagaaaaga	aaagtccatg	ggaacaaatc	ctcctcttaa	190860
gagcgtgaga	caggaatggg	gacacaagcc	aacaccccaa	ttgctaggct	aactctgata	190920
tgagacaaaa	gaatattaat	atcttggcta	tgaaggagga	tggtgccatc	ttctgaattg	190980
atgggagttt	tgaggcatgg	ctaagctggg	caaaccattt	tctttttttt	ctcttcttaa	191040
ttagtggttc	atztatggag	ggcttgcctg	cgggagagcc	catcagaaga	gagctcgctt	191100
tatggagatg	tagcttataa	aactactcag	attttaaaaca	aacagtgcag	gaggcagag	191160
gtagaagtgg	tgggggtggg	gtggggcaag	agaacaattg	catctgcaga	aggctagccc	191220
tgacacccaa	gcctatgttt	agggttgatc	agcttcccga	ggcaagccca	gaagcctcta	191280
aaatttttag	ccaatagaaa	tgacctctgc	accacggctg	actgaagcta	taaaataagcc	191340
tcgagttgag	cagtgggtgc	aacggagaga	gcagaggaaa	gtccaatcag	agcttcattt	191400
ttttttttta	aagtcacatt	gcttgggact	cacctgaagg	cagggcattg	agtagagcct	191460
tggctccctg	cagcgagagg	ctccagtttt	cccaggcacc	agcccatcgg	ttggttacct	191520
aaccaccgaa	agggaaactgc	acagcacaca	agttaaatat	aggctggggt	atctgcattt	191580
tacaagctct	gagcaagcta	tctgaagaag	ctgtcatttt	taatgacggc	acaaacttcc	191640
aattaccgac	tgggtaatcc	actagggagc	aggtagtttt	ggaagaacag	ttcaccatta	191700
ttaaaagttt	acacaatcac	ttttgagttg	actataagta	tttcacacga	ggcagggtggg	191760
attagggact	ttttgggtgg	tttactcgag	gctgcaacca	acaatgagtg	ttttctcaag	191820
aattatacat	tgagatttgt	caactgctgg	ggagtagtgg	agggtcctgg	taatgcagaa	191880
aggttatgaa	atggccaggt	aagggtgggt	gcttccaagt	ctcaaataata	ctcctaaggc	191940
cagctccaag	tcataagctc	aaacaagtct	tcaaggggcc	tggagagtta	agacaaaataa	192000
ggatcactta	ggctacccac	ggacaagcac	ttctcataca	aggaccggct	acctccaaca	192060
ccatcttccc	aacatggctt	ctatgttgct	tcaacaacca	gggcagggtg	aattaggggt	192120
gggtctctcc	aatgtggact	caaatcatga	ctacagcntg	gggttttttt	tttttttttt	192180
tttttttttt	tnnttggttt	ttcgagacag	ggtttctcca	tatagccctg	gctgtcctgg	192240
aactcacttt	gtagaccagg	ctggcctcgg	actcagaaac	ccgcctgcct	ctgcctctgc	192300
ctcctgagtg	ctggaattaa	agggtgatgc	taccacgccc	ggccgagtc	gtcttgataa	192360
tgaagttccc	agtgcacctg	atgtcaactg	aagttggatt	ttactgtgat	gactactgag	192420
tccggctcag	aattttgggg	ggacaaggta	actgggactt	actgggcact	acacgaggtg	192480
aacccccaca	ttgggagagg	cagaggcaga	ggcagaggca	gaaagttagt	tggaggctag	192540
gcaaggctac	acagcaagaa	gctgtctcaa	aaccaaagac	atctttcttg	atccaaatcc	192600
tgctcgaggg	tgtgaggcct	tggggggccag	aacaaggtgg	tcaagggaaga	ccactgactc	192660
tgctcctttg	tccattactt	aatcagaatc	gccatcacag	atatagctag	gagattttta	192720
gccttggtgg	ctgcaatctg	catttaagag	ctaagtggga	taaactcagg	ggtggggcca	192780
atgctccctc	ccccaccctc	cctgcacccc	tccattttac	tgtttccagg	gatctgctta	192840
atttacctgc	cagccttttg	tgggacacag	gcttagtggc	ttagcgctgc	tcggggcacc	192900
agagaccctc	acagaagcac	ctgaatgtac	tttcagcgct	gcagagcacg	cacggctcag	192960
gcccacacaga	agaacccagg	cttatgtctaa	ggagccagaa	agtagaagca	gctggcaaga	193020
gtgattcagc	cccataaatt	tacacatccg	tacagccaaa	cccacttgaa	gtgatccaga	193080
gccactttta	ttgaaaataga	aaagatgcct	attctggagt	gctaagtggg	acaggaggtg	193140
gggtatataa	gagataatcc	catgttgtct	ttgatgtggt	gctagggaga	taaccaggga	193200
cctcacgcct	gcctgcaagg	tagccaccaa	gccacaccca	caacctctat	ttatacacac	193260
actaagtgtg	gaggtatgga	taaaaaaaaa	tgtcccaaga	cctcacgaat	ctgcaaacat	193320
ggtgcctggg	tgggtggcacc	gtttggggag	gcagtggacc	atgttggtctt	gcaggaggaa	193380
gttatgtcac	tgggtatggg	ctttgagagt	ttgtagcttt	gctccccttc	cagtttaactc	193440
tgctctcgta	aggttcctgc	cacctgtttt	cctctgccat	tatggacacc	tggtcctcta	193500
gaactgtaag	ccacttactc	tcaggtctct	ttcagtcctg	gagtcttatc	atagcaatga	193560
aaagtaactt	gtgtggcagc	cagctaagca	agggtctggt	ccgactgctt	gggattatgg	193620
ttgtgtctgt	ctgtctgtct	gtcattccat	ttatatagtc	ctgagaattg	aaccacttta	193680
ccactgacat	gtctcagtc	tcttggtatc	atatattcac	ttaagacaag	atctctattaa	193740
gtctactcaga	cttggttttga	gcttgcaatc	ctcctgcctc	tgcttcaagg	cgataggatc	193800
cctaggggtac	tcgaccagac	tgggagtagc	aggttctggt	ctcttagctt	tctacagtga	193860
ttgtggatta	tttgtgtata	aagatctgat	ggcccagacc	actccccttc	ctttaagtga	193920
acatcaacag	tatttagcat	caacttaata	aactcatttg	gtaaagccat	ctccccacct	193980
cttgacaacaa	tgaaaatcaa	acagcagtac	ctgttctcct	agagcagcgg	ctctcagcct	194040
tccggccttt	taatacagtt	cctcgtgttg	cggtgacccc	ccccccccc	agccgtagaa	194100
ttatttcatt	gcttaaccag	agtttaactg	aagggttaat	aataaaacca	gtctggggaga	194160
ctaaggttac	ccaaccacgc	taggaaggag	aggaaagggc	cactcgcaca	aacctgtctt	194220
tgagatgaag	aacaatcaac	ataacaggga	cagagcagtc	cttgtaacaa	gtgcaaagga	194280

gagagagagg	ctgagtttct	acttctataa	ataaacctt	ggcaggcgga	tcactaaagg	194340
aacacaagtc	aatataaacc	tttagacatg	gggtcgccaa	acttcacttt	tcgacagtat	194400
attaattatg	tagtcaatag	ccatgggttt	cattagcgta	ttaaatacca	cgatcaatat	194460
tattttatact	tttcgaagac	aagccactca	gggaaaaaat	ggtgggggga	ggaggaggaa	194520
caatttgacc	ctgtagttca	aaaaaagtca	gaacagcaca	ctagagatta	gcaagggttt	194580
aatggaaggc	ataaaacact	ggaaatatgg	acagaaatca	gatccctgcc	ttcatttttc	194640
tgcccttttac	aaagagactg	gagggaaattc	agaaactatt	taaaaataaag	gcaaaatgat	194700
tagagcccct	ccctcccctc	agctgcttaa	cactgggggt	gtggtggacg	caaaataaagc	194760
attgagctct	aagtgataga	tgagaatcag	aacaggaaca	gtgtttttga	ggcaaaatat	194820
gtccaagaga	attcaaagaa	ctgtgggcca	gaatctactt	aggcagtcct	ctgggacccg	194880
aatccctcac	aggcggttaac	agtggaacca	atttccaagg	cagccctgct	ggtgatctga	194940
tttttgagta	gggaaatctg	ttaaacatcg	tcccacgagg	gagcccagct	ctttcactcc	195000
ccacgggttt	ctacatgcag	ctgtgctaga	tctgctgaag	tggccggtga	ggagggtgtg	195060
ggattgggttc	agcgacctca	gaggacattc	ttgttacta	gccctcgtgc	actggggcga	195120
tgaccgaatg	ctgtgagcag	gagatatcaa	aggccggcta	ctggactgaa	aactagatca	195180
ccatctctaa	cctgcaattt	gtcaatctca	gacagcaatg	aagactgtga	ttttctagt	195240
aacgctttgt	aagcaaggtc	agatagaggc	tccataaaaa	ttgttcagg	ttcaggcaga	195300
gaatcaagtg	taactcaatc	cctatctcct	gagatttagg	aagggaagga	aggctgtgtc	195360
tactaaacca	gtgagcctca	agcaaagcct	gtctgttctc	agcaagggtga	gccaccacc	195420
aaagatgcc	acagctaagg	gccagggatg	tagtgcagg	tgctgtgata	tcaacagctg	195480
ggagacagaa	acaggaggat	caggacttca	aggtagtttg	ggctataaaa	tataagcttg	195540
aagctaccca	cttgaagact	gtccccaaca	aaacaaaca	gctgggtatg	gtggtcgatg	195600
cttgatattc	tagtgtatga	gacgaaggaa	gaaagctcaa	gcccagagacc	tgcttgggtt	195660
acatagggaa	gatgtgtcct	caaaaaacagg	acagccgagg	agcagacaga	cagggcagac	195720
agggtgcac	gatctaaatc	cacatacctg	gatttaaagt	aatatctggg	agactggtct	195780
gtgaggccg	ttccagagat	ttaataaaga	cccaccctga	ctgagtatgg	gcaacacca	195840
tgggtggccc	aggggtccag	actgaataaa	gggaaaatgg	gaggaagttc	agctggtagt	195900
gtttccaagt	attaggacca	cagcctggcc	cctggcatgt	gctggccagc	tagtctagct	195960
ccagtatcaa	gcttcaggcc	acgpgcagg	cactggacag	ttcccacaca	cgacacacac	196020
acacacagag	cactagcatt	cacctcctgg	tctcttcttg	acaacagata	aaatgtaact	196080
ggctgccaca	gtgagagtcc	cctaccttcc	tcaccgttaa	ggatggtaca	aactgtgagc	196140
cagcagcagc	catttctcca	gtaacttget	ttccacagat	actgttatag	cactaagaaa	196200
agcaactgaa	acatggggtg	ctgtgacccc	ttggcaccac	aaagccatgg	caagctgaag	196260
tgacacatc	acaggccagg	cctgaagatg	ctgggggact	gcaatgctgc	ctggattctg	196320
gcagagatgt	gcagcagatg	ccaagagggtg	ggctgcagca	accagagata	attaatatga	196380
ttaggaacac	actgagcagg	catgctcttg	ccgaatgaaa	agcctcgag	tgtaatgact	196440
gttttcttcc	tcgatcacgg	tctccacgtt	tcagaagttg	cttggtgtta	ggctgcccgg	196500
taaaacatcaa	tccaacccc	aggggcccaga	tcacggtgt	tcctgggctc	aatcgccctt	196560
ccttttgtgt	tttcattcat	ttaaagatgc	attccagggt	tgcaaacatt	agtgagaatc	196620
atctccaggc	ctcgtctaa	tctctgagtc	tgtaatgagt	taacatcttt	ccctagtga	196680
tatttattat	gaaggcta	taattgcttt	ccagttaca	gaatccttta	cagtcaaaga	196740
aagtaggatc	cacaaagata	tactgtttat	tcaaacaaag	caaaggaaac	aaagcttctt	196800
tcttaaattc	tatttaacat	agctttaata	aaggtaacac	ggtccgcctg	gcaaccgaac	196860
ggtaactgat	gcaaactgaa	gccatgctct	gtagcagcct	ggatgtccca	gtgccacctc	196920
tgctgcagg	ctttgtcgga	tttactaaga	ttctgttata	ttcaaacagg	gattgtgtct	196980
caagtaactg	accccactat	gtggataatg	aagtaaatta	tgcaatttgg	gggtttgctt	197040
ttccccaagg	ggacagcaag	ccagtgttta	tcagccgtcc	tcagaggaga	caattctgat	197100
taatatcaga	gtcatctgac	tcagtctatt	aaacctatca	aacctgaag	gaaggatatt	197160
cagatattaa	cgataggcct	ttgattaata	attctacctt	gttgccattc	taagcattaa	197220
caaccatgca	gtaactctgc	aaaacagacc	ccttgattcc	aggcagacgc	accctctgaa	197280
cacctgggtt	ctcccctact	cttctcccc	caggaggaac	tcaagacaaa	aagggtccac	197340
cactggaaaa	gcacactcca	ggttacataa	tttgccctcat	tatccagagt	gggtttaatg	197400
acttgtgaca	taatttctgt	ttgaagataa	caaaatttca	tgaaatccga	caaagccgga	197460
aggcaggagg	aggggactgc	tgccacacta	ccggtggctg	agaactggag	cggaagggtt	197520
acacacagcc	ctctgagctc	actgtctttg	cttatcagtg	agtcacca	ggggcccaga	197580
tgggttgcca	gcctccccta	gaggatcttc	attgtggagc	tgtcccatgg	ggcgggaagg	197640
aagccattct	atcttctgtt	ttctctcttc	cgttctggcc	accagtggta	cttgctccca	197700
tcacatgttc	ttcctgatgt	tcgcatcag	ccgtctgcca	tagtctctga	agtccacggg	197760
cttcacgtcc	atcacagtgg	ccttaattcg	agattcatcc	tttagaaaag	agagaagctg	197820
tttgtgagtg	gcagagcctg	gcgtgcagcg	gaagagagaa	ctttctttgc	ttcagtggct	197880
tcaatgagtc	cagcaggaag	aaggaaagtt	tacaagtctc	agagagaaa	tgctgtgact	197940
tcctggagtt	gggcccagtc	ctcttcacac	gaccttttcc	ccatccatagg	tgccctgtgc	198000
tcagacctag	catcctcccc	gagaagcctc	tgtctttcta	tgggtgcagt	gggggcccag	198060
agcagacagg	taactcacc	taaagcatca	ctttcatcta	gaggagctct	gtggtagtag	198120

ggactgaggc	ttctgctcca	gctctgggca	aggttacttc	tctgctcttc	accatttcctg	198180
tccatcccag	gaagacagaa	aatccctaca	ctctcccttg	atctacccga	ctttctgaca	198240
ccagcctacc	tatgttcatt	taatacaaca	actaaaatat	ctattcacag	gcactaagct	198300
ggtgataacg	cagaatgcac	aaactctgcg	gctgcagggg	agacggcaga	gttcctcctc	198360
cacttgcttc	cttgaactaa	acagtgtcct	tgaggcagaa	cagggtgaca	cctagggaca	198420
cacaagtcta	gctggggggc	ttcatgcttc	catgtgctta	gtaattaatt	actacatgca	198480
ccgctgttta	caagtatggt	taggagcccg	actgcctggg	ttggcctctc	gcctctgcca	198540
ctccatggct	ttaggttcag	agtcattctc	tgcatgcctc	tgctgtcttc	tcggttggtta	198600
aagcttgcaa	caacagctcc	aacacagaaa	gtgctgtgag	ggtcgacagt	ggatagatgg	198660
ctagatagat	ggggcaggac	ggactgtcca	gtaagcaggg	ttcatcatgg	ctatgcagct	198720
ctggacatca	ggattagttt	aaacacttgt	cagggtgggc	actttttacca	gcacgtgcta	198780
tttgtttaat	attctgagtt	ttagaaccta	aactgtggga	aacaagagtc	cacacataac	198840
annnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	198900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	198960
ctcctgctgg	ggttacaggc	ttgagccacc	atgacagctt	tagcaatagc	tttgtaaatac	199020
cacagtgtca	agctggatat	gatggcacat	gcttgcaata	ctaacctcca	aagattccct	199080
gaactcgagt	tggtgaaata	gtccaccagg	taaaagagct	tgctgcccac	gcctgaatct	199140
gattccctgg	tcacatgctt	taaggaaaga	acttgccgaa	gatgtcctct	gagtgccacg	199200
tgtaccgatg	catgcttgca	gccacccaca	caccacacac	agtgcactgt	ctcacacagt	199260
gagaacagca	agtgaacaaa	caaacaagcc	gggggggggg	ggattgtgac	cagaataatt	199320
gagggggggt	gtaaagctct	tggcaggtgg	ctggctcctg	gtaacactcc	ataagtgggg	199380
aagttccaca	tgtaaggtca	tgtgatcgag	tacatctggg	cctccaacag	tccttgnnga	199440
agaaacagat	gcagctctgt	atatctaaac	catgttgttc	gtatctctgg	gtagtctttc	199500
ttttctcctt	cctttctttt	ttctctccct	ttctctttta	aaaaattatt	tattttattat	199560
tatatctaag	tacactgtag	ctgtcttcaa	acacaccaga	agaggggtgt	tgctcagatct	199620
cattatggat	ggttgtgagc	catcatgtgg	ttgctgggat	ttgaactcag	gaccttcaga	199680
agaacagtca	gtgctcttaa	ccgctgagcc	atctctccat	cccccaaccc	ctttctcttt	199740
tgagttaggt	tttgtgtagc	cctgggtggc	gttaccttaa	ctacactggc	tttgaacttg	199800
caatgatact	ctgcctgata	tgtcttaata	attttgagat	agggactcac	tacatagcct	199860
ttgctggtct	ggaactaaca	gagatctgcc	tgtttctgcc	ttgcaaatgc	tggaataaaa	199920
gttatgtacc	accacacctg	gagtttaagg	gttttttggt	tgtttggttt	tcgagacagg	199980
gtttctctgg	gtagtcctgg	ctgtcctgga	actcgctctg	tagaccaggc	tggtcctgaa	200040
ctcagaaatc	cgctgcctc	tgcctcccaa	gtgctgggat	taaaggcgtg	tgccaccacg	200100
cccagtttaa	gggttttttt	ttgtttggtt	tttctgaga	cagggtttct	ctgtgtagct	200160
ctggtgttct	tggaactcac	tctgtagagc	aggttagcct	tgaactcaga	aatctgactg	200220
cctctgcctc	ccaagtgtcg	ggattaaagg	cgtgtgccat	cactgcccag	tgattttttt	200280
ttttttaatg	tgtgtatttg	tatgggtgtg	tggtgtcttg	tggaagccag	gtgtcagatc	200340
cccagagcgg	aagtgttctt	aaccgctgaa	ccatctctct	tccctcttcc	ctaactctga	200400
ttttaaaagg	accaaactct	taggttaggag	actatacaca	cacacacaca	cacacacaca	200460
cacaccctga	cacaccctga	gacctgcctc	gacctgcctc	gagcacacaa	gtggttttat	200520
tgctggtctg	gcctgtgtat	gagctggaac	caaaaccttt	gtcgggagat	ccgcagtctg	200580
cagtttgagc	acaggctctc	tggtttctgt	tctctgtcct	gtgtcgcata	ttgactagag	200640
gcagagaagc	atctgcaagg	ctgtgaccac	gctggctggt	gctctgccat	ctacattttc	200700
aacaggaaat	ctcaggagag	tatttctctt	taagaacgcc	agacttttgt	gcctgggcca	200760
cttctctact	tcacagaaca	ttgtgtgcca	agtggcaagt	tattaaccaa	gtgctttgga	200820
aaattaaaact	ccttggtttg	cagagtagca	tgaggagcatt	gagaggggtg	atgcctaaag	200880
gcctggttct	gctgctggca	gagctgacac	ttggctaaag	ggctggcatt	tctgagatga	200940
gcctcactag	atccgcgtct	cagagtctgc	aggagaaatc	agagagggga	gaaggtccag	201000
tggtcctgtt	aggatgatct	tctctgcat	ttaagggcgg	ctggtttgcc	cacgtagccc	201060
cagaacaaaa	cgagcctcgg	acgaagcccc	ctaaaggcag	taggagagac	tgagccttgg	201120
ctcttcagca	gggttgggga	caagagcaag	aggcgggatc	tcgcccggcc	ctttagagac	201180
acgtgcggtt	gtttccgtgt	ctgggagatc	acatgaccgg	catcagctga	cccgtcacgg	201240
tgagctcag	cgctggtgct	tcgcgtctcc	cgccctgctg	cgccccggag	cgcaggaccc	201300
tgcgaggagg	taagaaaacc	cccaggcttt	ctttcctttg	tcgctggttc	gcgcagtcac	201360
ctgcaccccta	ccccccgctc	ctcgttcata	ccagtcttcc	cggcctggca	ccccggaagc	201420
cactgcgagg	agggccgtgg	ccaggctcag	ccttgctgct	ccccaggcgg	gccaggacca	201480
aatggcccaag	gggagcagaa	ggcggaaggt	ggttcttaca	gcaggggtccg	agggtggttc	201540
cccttctctca	ggacctgaca	tgaggagact	gctccggagc	gtggagagag	atctgaacat	201600
tgatgcccg	cagctggccc	tggcgcgggg	gggcactcat	gtagtggccc	tagtgtccac	201660
gcgttggtcg	gctagtctcc	gggagcgccg	actgggaccc	tgtccccggg	ctgagggcct	201720
gggtgaagca	gaagtcagga	ctttactgca	acgttcggta	cagaggctgc	ccccaggctg	201780
gactcagatg	gaggtgcatg	ggctgcggaa	acggagactg	tcctacccgc	tgggtggagg	201840
cgtgcccttt	gaggaggggt	cctgtagccc	tgaaactctc	actcggttca	tgacaggagt	201900
ggctgcccag	aattaccgga	acctgtggcg	ccatgcatac	cacacttatg	gacagcctta	201960

cagccacagc	actgccccct	cagctctacc	tgccctagac	tctatacgac	aagctctcca	202020
gaggggtgat	ggatgcacct	tcttgccagt	gggtgaatcc	atcccatgtc	tatcaaagt	202080
cagggatggg	ccctgcccc	ctcggggcag	ccctgctgc	cccagccttt	tgcgagctga	202140
ggctttgctg	gagtcgccc	agatgctcta	tgtggtacac	ccttatgtgc	aattctccct	202200
gcatgatgta	gttaccttca	gccctgcaa	gctgaccaac	agccaagcca	aggtgctctt	202260
tcttctcttc	cgtgttctga	gggccatgga	tgccgtgcac	cgccaggggc	tgccctgtgg	202320
ggctctgtct	ttgcaccaca	ttgctgtaga	cgagaagcta	tgcatgtgagc	tccggctgga	202380
cctgagcgct	tacgagatgc	cttccgagga	tgaaaaccag	gagggctctg	aagagaaaaa	202440
tgggacaggc	attaagtctg	aaaaagaggg	ggaaggagga	actgagtgtc	ccacctgccca	202500
gaaagaactt	cggggccttg	tgctagactg	ggtccatggc	cgaatcagca	acttccacta	202560
cctcatgcag	ctgaatcggg	tggcaggtcg	acggcagggg	gatcccaact	atcacccagt	202620
gctgcctctg	gtgggtggact	ttaccacacc	ttatgggcgc	ttccgagacc	ttcgtaaatac	202680
caagttccga	ctcaacaagg	gagataagca	attggacttc	acctatgaga	tgaccgggca	202740
ggcatttgtt	gcaggtggtg	caggaagtgg	ggagccaccc	catgttccctc	accacatctc	202800
tgacgtgctc	tctgacatca	cgtactatgt	atacaaggcc	cgtcgcacac	cgcgctcggg	202860
gctctgtgga	catgtccgag	cgcatgtgga	accccacgag	tatcctgcca	ccatggagcg	202920
gatgcagacc	tggaacccgg	atgagtgcac	acccgagttc	tacacggacc	cctctatctt	202980
ttgctctatc	caccttgaca	tgcccgcact	ggatgtgccg	gcttggtgca	gttctaacca	203040
ggaatttgtg	gctgcccatac	gagccctcct	ggagagctgg	gaggtgtccc	aagacctgca	203100
tactgggatt	gatcttacct	ttggctacaa	actccagggc	aaagaagctg	tgaaggagaa	203160
gaatgtgtgt	ctgcacctgg	tggaagctca	cacctatctg	accagctatg	gcgtggtaca	203220
gctatttgat	cagccacacc	cccaacgcct	ggctggatct	cctgccctgg	cccctgaacc	203280
tccactcatc	ccccggctgt	tggtccagcc	tattcgggag	gccacaggcc	aggaggacat	203340
ttcaggacaa	cttataaatg	gtgcggggcag	gcttgtcgta	gaggccactc	catgtgagac	203400
tggtggaact	agagataggc	ctgggacagg	agaagatgat	ttagaacagg	ctacagaagc	203460
tctggattcc	atctccctcc	ccgggaaagc	aggtgaccag	ccaggctctt	cctccagtca	203520
agcatcacct	ggcctgttgt	ctttttctgc	accctcgggg	tctcgaccag	gccgtaggag	203580
caaagtgcgc	gggttggaac	ctggggaggg	tgaagagggc	aagattgtcc	ttccagaggg	203640
cttcagtcctc	atacaggcct	tggaagagct	ggagaaagtg	ggtaacttcc	tgcccaaagg	203700
cctagggagc	cagttggagg	agcctgaaaa	gcctcacgcc	cagccacctg	tgccacctgca	203760
gagcctcttc	catcgagaca	tgccaggtcct	gggtgtcctg	ttggctgaga	tggtgtttgc	203820
caccaggggc	cggtactgac	agcctgatgc	acctttgtgg	gtacgctttg	aggctgttcg	203880
gggtctctgc	atacggcact	ccaaggacat	ccccgtgtct	ctgcagcctg	tgtagacac	203940
actcctacag	ctagcggac	ccaaaagtcc	catggtgtcg	aagaagggca	agctagaccc	204000
actgtttgag	tataggccgg	tttccagggg	attaccccca	cccagcccag	cccagctcct	204060
cagcccttcc	agctccgtgg	tccccttccc	tccatacttc	ccagcactgc	acaagttcat	204120
tcttttatat	caggcccggc	gtgtggagga	tgaggtccag	ggtcggggagc	tggtgtttgc	204180
tctgtggcag	cagctgggtg	cggtgttaaa	tgacatcact	cccaggggct	tagagatcct	204240
cctgcctttc	gtgctgtcgc	tcatgtctga	cgccacacag	gctgtgtaca	cagcctggta	204300
cctatttgaa	cccgttgcca	aggccctggg	ccccaaaaat	gccaacaagt	acctcctgaa	204360
gcctctcatc	ggtgcctatg	agagcccctg	ccgcctgcac	ggccgcttct	acctgtacac	204420
cgactgtttt	gtggcccagt	tggtggtgcg	gctgggcttg	caggccttcc	tcacccacct	204480
gctgccccat	gtcctccagg	tactggctgg	ggtggaggct	tcccaggagg	agggcaaagg	204540
cctggtcggg	accactgagg	atgaggaaa	tgagctcccg	gtgtccgggc	ctggctcctg	204600
tgcccttggg	gaagagattc	agatggatgg	gcagccggct	gcttcctcag	gactggggct	204660
cccagactac	aggtcggggc	tcagcttcca	tgaccaggcc	gacctgccgg	acacggagga	204720
cttccaagct	ggactctacg	tggtgaatc	tccacagccc	caggaggctg	aggccgtgag	204780
cctggggcag	ctgagtgata	agagcagtac	cagcgaagcc	tcccaggggc	aggagagggg	204840
tggggatgat	ggcgggtgcc	ctgcggacaa	gaacagcgtc	aagtcagggg	acagcagcca	204900
ggacttgaag	cagagcgaa	gctctgagga	agaggaggag	gaggaaggct	gtgtggtgtt	204960
ggaggaggac	caggaggatg	aagtcacggg	aacatccgag	ctcactctgt	ctgacacgat	205020
gctgtccatg	gagacggtgg	tggtccctgg	tgatgggaga	gacagagaag	aggaagagga	205080
gccgctgaca	gagcagacag	aaggcaaaaga	acaaaagatc	ctccttgggtg	agcccgtggg	205140
ctgagggggc	atgggtcagg	tgcttttctc	tcaggctctc	atatgctggg	tgtgggtcca	205200
accagatcca	ctgtagcacg	cacagccaca	gtgcatggaa	gtgcatggaa	tgtgggaagtg	205260
ctgtgtgtga	gtggaaaagtg	gggcttagat	ttagctttca	ggagacagaa	agctccttta	205320
aaagccatac	cttgggctga	ggctgggagt	ggagttagtg	ggtagagcac	ttgtctggta	205380
tactcgaggg	ccctgggtgt	cttatctcta	gccccagaag	aagtattaa	aaataaaaagc	205440
aagtgggtgt	tgagatgtga	atggagccag	aactggccgg	aacagtcggg	tggaagtggg	205500
agagtggttc	cagacagggg	acagtgtgtg	tgtacctctg	aggctctcat	ggttccatca	205560
gagagtcagg	gaaaggctaa	aatgggtttc	ttaaagagtg	ccagaagggc	tggtgctggg	205620
ggtgcaggcc	tttaatccca	gcactcagga	ggcagaggca	ggcggatttt	tgagttcgag	205680
gccagcctgg	tctacaaaagt	gagttccagg	acagccaggg	ctatacagag	aaaccctgtc	205740
tcgaaaaaaa	aaaaaaaaaa	aaagagtgtg	gaagggtgga	agccagggac	aagtctgtac	205800

aagaaggaac	ttgggagcat	tgccgaaagg	atgacctctc	tgcaggctcct	gcccagaggag	205860
ccagttttctg	gggaccttga	ccatggctag	gtgaatggac	ccaggatggg	atgggtcaggc	205920
ttgctagcag	agccacagcc	gagttggctg	ggtgggggtg	ggtgggggtg	gaaggggtgag	205980
ttatctgatg	agctcaggac	cttttccctg	cctgcagata	cagcctgcaa	gatgggtccgc	206040
tggtctgtctg	ccaagcttgg	ccccacagta	gcctctcgcc	atgtggcccg	gaacctgctg	206100
cgctgtctga	catcttgcta	tggttgtaag	gtctgtgggt	agtgtctggag	accaggttcc	206160
ccagccaggc	ttctgcccac	ccttagccct	ctctaggcga	ctccttccct	aacttcccag	206220
cactccctga	gcagggcctg	ggtctcacc	attaagctgg	gttttcttgg	gtaagtgggg	206280
aagagcccag	tattgaatga	atagaagcca	ccccacagtc	tcagaaggcc	ggcttccctc	206340
ctgccctcca	ctggcttctc	aacgctgtg	cccttccctg	gtagggccca	ctcgacagca	206400
gttcaccgtc	agcagtgtg	acacccctcc	actgaatgcc	ggcaacatct	accagaagag	206460
gccagtccca	ggtgacatcg	tgctggggcc	tggtctcagc	tgccctccct	acattgccta	206520
cctgtatgga	gaacccgttc	tcacctacca	gtacctgccc	tacatcagct	acctggctcag	206580
tccttggttc	gtcaaacc	ggcttggggg	tgggggcaag	gatccaagga	ccagccccag	206640
gtcttggggg	ttccaggagg	tctgtgggg	gacctgtccc	tcctcatct	attctgtggt	206700
tctaggtagc	cccaggaggc	aactcaaacc	ccagccgact	gaacagccgc	aaggaggccg	206760
ggctgtctgg	agcggtgaca	ctgacgcaga	aatcatcgt	atacctctct	gacacgacc	206820
tcattggacat	tctgcccgc	attagccacg	aggctctgct	gcctgtgctt	ggcttccctca	206880
cctccttctg	cacagggtag	gcccctgctg	cttgggagag	ccacctggct	gagggggccc	206940
ccaggaaggg	ctaggaagct	caggagagaag	cagataccgg	cctgagtcac	ggttctgatg	207000
ttgggggttag	tgccacaggt	ctttcattcc	agcaccacaga	ggagggcaag	tttctgtgag	207060
tctgagacta	gcctggtcta	cagagagagc	tccaggctat	ctaaggctcc	atagtaagac	207120
tctgacttaa	gaaaagagtc	gtgggttcatt	ctgggttggtg	ggtgtggctt	ggtgatggga	207180
cactttccca	gcatgcagga	ggagctatgc	ttgagttcca	gcccttcaga	aaaacaaaaa	207240
tgggggctgg	aaagaatagc	tcagggttta	agagcactgg	ttgctcttcc	agaggatcca	207300
ggtttagattc	ccagctgcca	catggtagct	cataaccatc	cggcagttct	atggaacctg	207360
ccacccctct	tcggctctctg	tgggcactgc	aaacatgtgc	acagacatac	atgcaggcag	207420
aaaaaacacc	catacacata	aaattagacc	aaaaaagttc	atgttctctc	ctacctgtag	207480
ctctgactaa	gctacactgc	ttccctgtgc	ctcagtttcc	tcccctggtc	tggaactgatc	207540
agccttacat	gcagctcctg	ttatttgaag	ttcctggtaa	attgggtcaag	tccttcaggg	207600
aagggtctgg	aactcttgca	ctttgattct	aggttcccca	gtggggccca	ggcccggact	207660
gtcctatgcg	tgaaaacccat	cagtctcatc	gccctcatct	gcttgcgcat	cgggcaggag	207720
atggtccagc	agcacctgag	tgagccagtg	gccaccttct	tccaagtctt	ctctcatctg	207780
catgagcttc	ggcagcaggt	aggcaggcag	cttctgggct	gggtggggcca	ggccaggcca	207840
ggccaggcca	gggcagtgga	ccactgaat	ctgtggtctt	cctaccgcga	ggatctgcca	207900
ctggatccca	agggctgtac	tgagggccag	ctgccagagg	cgaccttctc	tgatgggcag	207960
cgacgaccag	tggaacccac	cctgctggaa	gagctgcaga	aggtgttcac	cctggaaatg	208020
gcgtacacaa	tctacgtacc	tttctcctgc	ctgttgggta	ttgcccatac	cgttcccttg	208080
cacagagttg	gtgactacat	ctcttccctg	gggtggggcc	cgatgctttc	acctccagag	208140
tcagcaatgg	aatcttttta	tttttatttt	gacatgggg	ctcatttagc	ccaggctgac	208200
ctttaactcc	agctccttcc	agcttccacc	gtctcctggt	ggcattgtag	tcatgtggca	208260
ttgctcaggc	ttcttncatg	ttcttatttt	taaatgacct	gtgtgtgtgt	gtgatattctg	208320
tgtagtgtg	gaggtgacag	aataacagtt	gggggggtcag	cagatgcctt	gcctgatgag	208380
cattctctca	gatccagttt	ttgggtttgt	gggcttttat	gtgtgtgtgt	gtttgtctgt	208440
ttttgtagac	agggtctctc	tggttagcct	ggccatccctg	gaactcattc	agtagaccac	208500
gctggccttg	agctcacaga	gattcacctg	cctctgcctc	ccagtgtctg	gattaaaggc	208560
gtgtaccact	cctgcctggc	tttgtttttg	ttaaccacca	tcctcctgcc	tcagcatctg	208620
cctccctctg	gctgggatta	caggtgtgtg	ctatcacacc	cagctaacag	tggaatttaa	208680
cgtaggaatt	ttaggatcag	agtgaccaga	tttggctcta	gggcccatt	tccacagtga	208740
ttatctatct	tagttaggat	ctctgttgaa	aatcatggtg	gaacatcatt	accaaatgca	208800
acttggggag	gaaaagggtt	attttgtctg	acaactctca	ggtcaccaag	ggaagtcagg	208860
gcaggaactc	gaggcagaag	ctgaagcaaa	agccatggaa	gaactctggc	ttgttccctca	208920
tggtctgtct	agtctggtgt	acccctctcc	caccacctc	cccacaatgg	tttctctgct	208980
tatgcctggc	tgctctagaa	ctcactctgt	agactaggct	ggcctcaaac	tcaagagatc	209040
cccctgcctc	tgcatctcaa	gtgctaagat	taaaggcggt	tgccatcacc	cctgccccag	209100
gggtggcact	accactgta	aattggctcc	ttcccatac	agttgttaaa	taagaaaact	209160
cctccatagg	ccaatctggt	gggggaattt	tctcagttga	gggtttctct	tctcaaataa	209220
ctgtagctga	tgccaaattg	ataaaacaaa	tctcaaacca	ccaccaccaa	caacaataaa	209280
accaaacaaa	ccaaacaact	aaccaagaca	gtgacttata	aagagaatct	gaacattttc	209340
cagcaggaaa	ggctcaggag	ctggccattc	aagctctggg	aacagaatgt	aggggaatat	209400
gatggtctcc	agaagctacc	tgcaaaggaa	tgcaacgctt	gctgggtttt	gtgggttccc	209460
ttatgggatg	ggcgctgtac	tgggcttctc	tctgagtagg	atgggccacc	ctgtagttgg	209520
gaatatattt	ctcctacaga	attgtaagtt	cccagaggca	ggacacatct	gtcttattct	209580
tcattgtgtg	tctgatgcta	gaatggtgcc	tggcatacac	gtgtgtgtct	ctatagagac	209640

agcactcatg	tctacgtatc	gataaaggaa	gctgttttgg	ggggaggaaa	caggcttaca	209700
gacgagaact	taataaccca	gagtagccca	gtcagtacct	tgccttggct	tctgtttgtt	209760
ctaagctctg	ggtagatagc	taccttgcca	tcttccttga	tcttagaact	ttccccactc	209820
ccctgtaggt	gacatcatcc	ggaaaatcat	ccccaaccat	gagttggctg	gggagctggc	209880
agggctctat	ctggaaagca	tgagcccgag	ctctcgaaac	ccagccagca	tggaaccac	209940
catggctagt	gccggccctg	aatgggaccc	tcagagtggg	agctgtctcc	aggacgatgg	210000
ccactcaggg	acctttggga	gtgtcctggg	tggaatcg	atccagatcc	ctgactctca	210060
gccccagagt	cctggggccac	tgggctccct	ctctggagtg	ggtagtagcg	gaggcctcag	210120
caacaggaat	gaagacaacg	ccctgaagcg	ggagctgcct	cggagtgcgc	atgggctgag	210180
cgggaactgg	ctggcgtact	ggcagtacga	gatcggtgtg	agccagcagg	atgcccactt	210240
ccacttccac	cagatccgcc	tgcaagcgtt	cccagggcac	acgggggccc	tcaaatgcgt	210300
ggcgcgcctg	agcagtgaag	acttctttct	gagtggcagc	aaggaccgga	ctgtgcgcct	210360
ctggccgctg	tacaactatg	gggacgggac	caatgagacg	gcttcccgcc	tcatctatgc	210420
ccagaccgcg	aaaagcgtct	tctacgtggg	ccagcttgag	gccccgcagt	atgtggtagg	210480
ctgtgatggg	gcagtgcacg	tctgggaccc	cttcacaggt	gagcggggcc	aggtgaggcc	210540
gtttcgacgg	ctgctttact	gtgccttagc	caggcctctg	ggaacgggac	ctagtgcgaa	210600
acgtacaatg	gcgtattttg	acggggaaga	ttcagtgaag	caggaagaga	agaagagtca	210660
ggacttagaa	tctgtgggac	ccaagtttga	atccactccc	ccaacttacc	agcaatcggc	210720
tcagttgctg	caggcgtctg	ccttctacct	gtaagaacca	aaaattttaga	agattccacg	210780
agtatggctt	tggcttcttg	tacgacgtca	cctgtcgtcg	ttgtaaagag	aagtatcgag	210840
tggaggaggg	tcagggcaga	cggaggtcgc	agctagttag	agcatgctat	gtgaagagag	210900
cagactgttc	tggggctgga	cccttgactt	cactgtggaa	gcagcaagat	gagaaagccc	210960
tgagattgtg	ttttctgagg	gtcactgggg	aatgggatgc	aggtgtgggg	tgagttggag	211020
tttgaagtag	ccagggctct	ttgatagcca	ctaagtcccc	agatgtgtcc	tttttcagga	211080
aagacccttc	gcacagtgga	tccttcagac	agccgggtgc	ccctgacggc	tgtggctgtc	211140
atgcctgccc	cacacaccag	catcaccatg	gccagctccg	actccactct	gcgctttgtg	211200
gactgcagga	agccaggctt	gcaggctcag	aggggtgcag	ttcctgggct	actgggggtc	211260
tctaggtacc	agtcaggaaa	gacactcagg	ggactccacc	aggaacgctg	cagtgcacag	211320
cagccctgtg	tgggtggggc	gctggcacgg	atggggcttt	tctcttccgg	ggatggagtg	211380
ggagggtcag	gcctactggt	ttcgtggggc	tgaatggggg	gagctgcagt	aggggtgggtg	211440
gcagtgatgg	atggcgacgg	gcacttgaa	acaatctcct	cctatagcat	gagttccgac	211500
tgggtggagg	gctgaaccct	gggcttgttc	gctcgttggc	cgtcagcccc	agtggccgga	211560
gtgttgggc	tggcttctcc	tcgggcttca	tggtgtcctc	agataaccgc	acgggctgtg	211620
ttctacgagg	ctggccagcc	catgaagggg	acttctaca	gatcaagggtg	actgactgcc	211680
tgaggtccta	tcctttcatt	tctacttagg	gcctggctctg	ggagaggaca	ggtttatgct	211740
ggtgtccctt	ataactactc	ggggacattc	agtggggtgg	gaaaatggcc	ctcgtaggcc	211800
agctcaggaa	ccagctgcac	aggaggcagg	ctaggggcag	gaatcagggc	tagaactgac	211860
cctgatgctc	cacagcgatg	ttctaattag	taacccttgt	ccataattgt	cttgcttggg	211920
ggatcagggg	tcacgccctg	tccgtgacct	agttcaggtt	aaataaagcc	aggaggctgt	211980
ttactgcctg	gagaccactg	agcagagtcc	atgccccctg	ctgggctgtc	ctgatggggg	212040
gcaggaacag	gcgcaggcct	gcgcctcgtg	ttcctgcctc	ctatattcaa	tcatagacct	212100
cagagctcag	caggttctgg	gaggggagaa	atagggctgc	ttgtgggagg	atcttctccct	212160
gcagtgggaa	ctctcctccc	ccgccgtcca	gatggaggtg	aaagacaggc	actgttgcctt	212220
acaggaagg	caggctgccc	ccagctctat	ccaggacccc	aggggaccct	gggtctcagt	212280
gtctctaaat	cccaacattc	taagaaagt	tcaggatggc	tctggggtca	tcctgggtgt	212340
tagtcccagc	tctcggagtc	ttctctgagc	accagtctct	ttcctatggg	aagtaaggac	212400
atgccaggtg	ttctttgaga	ggacctgagt	ttggttctca	gcactgtcta	gctctggctc	212460
cagggggttc	aacacccttt	atggcttctg	tggaacacata	ttcttatgtg	gcctgcacgc	212520
acacacacga	acacaaataa	aaataaatgt	taaaagaaga	cagcggcacc	ttgtacctca	212580
catgttagta	cgattggatg	tggcagtgcc	tcacagaatc	tgtgggactt	tatttattta	212640
tttatttttt	ggtcttaaaa	ttttaaaaga	ttggtttagg	agtgggtgga	cacaccatta	212700
atcgcaacac	tcaggagcag	aggcaggtgg	atctctatgg	gtttgaggcc	agcctggctc	212760
acagagcaag	tttcagggca	gccaaagtta	cacagagaaa	ctctttctca	aaaaataaaa	212820
acaaaaatatt	taaaagattt	acttacttct	tatttgagct	aggatctccc	tattagccct	212880
ggctgtcctg	gaactcactg	tatagaccag	gctggcacct	taaactcacg	aagatccctc	212940
tgcttctgcc	tcctaagtgc	tgagattaaa	gtagtgttat	accatgcccc	actattttct	213000
ttatagattt	ggtgttttgc	ctgcttgtgt	atatatgcac	taccttcacg	cagtgcataa	213060
aggggtcaga	ggcgagtatc	agctcttctc	ggaactagag	ttatggaagg	ttgggagtca	213120
ccatgtgtgg	actgggtcat	ttgcaagagt	tacaagtact	tctgagccat	ctccagcccc	213180
ctagagtttt	tttcccccct	ggctgtcctg	aagtagaatc	tgttcttgtt	ttgttttgtt	213240
tttcagagaca	gggttttctc	acataggcct	ggctgtcctg	gaactcactc	tgtagaccag	213300
gctggcctcg	aactcagaaa	tcgcctgcgc	tctggctctc	agaatactgg	aattaaaggt	213360
gtgcgcacc	acgcctggct	cagaatctgt	ttttaaatga	gagtaaatgt	tacaggtttt	213420
ttgttttgtt	tttttctttt	ctttttttgt	ttttgttttt	tggctcattt	gttttatttg	213480

```

ttttgagaca ggtctcactc tgaaccctta ggtggcctgg agcttgctat gtagaacaca 213540
ctgactttta acttgttttc tgagtgtctg atttatgggc ttgtgctatt ttgcccagcc 213600
tctgatgggt gtttaataaca atattattta gcttttcttt tggagatagg ctctcactgt 213660
atatcaccca gacagtggct ggtctggaaa tcaactgtgtta ggccaggctg accttgaatt 213720
cacagagatc tgcctcccga gttcagaaat taaaagcact ctgggatgggt tttggagttt 213780
ggtgagtacc caagcctcca ttgatgctat ctgtccctcc cgctctctgc aggctgtaga 213840
gggcagcgtg ctcatcagct cctcttccga ccattccttg actgtttgga aggagctgga 213900
acagaagccc acgcaccact acaagtcagc gtccgaccca atccacacct ttgacctgta 213960
cggcagcgag gtggtcaccg gcaactgtagc caacaagatt ggtgtctgtt ccctgcttga 214020
gccaccctct caggccacca caaagctcag ttccgagaac ttccgtggca cgctcactag 214080
tctggctttg ctgcccacga aacgccacct cctgctgggc tcggacaatg gcatcatccg 214140
cctcctggca tagggccagc caggagttag gctagggcag ggcgagatga catctctcag 214200
ggcccgtctc tcattcttga tctcgaagcc gattcttcta ggcaagcccc aggctctggc 214260
taccacatg gctgtctgtc tgggattgca cagctcctga atctccaaag ccttgaagtg 214320
gcttcatgaa actcgggaga tactgttccct aaccagcaag aattggggca aggaaagcac 214380
tgtgatcccc attgctcccc agttctgcct tctggattca catggggaca gggcagctcc 214440
aggaaatgaa aggagtggg cctttgctca gccagcttcc tctagccacg ctctccttag 214500
ctctgtttct cccttgggta ggaaactgct cctgtctagg gttctgatgg tactgggact 214560
ccaggctcag gagggtggc caggacctac gactttcagg gcttgggtctg gggtttttagc 214620
attcattcag ccaggctctc agtatgggac cagaaaaaag gggatgtgag aacagggtcta 214680
gggaaggggt tatatgggcc cagctgggtcc aggaatgaat ccatgccttg ccttgggtacc 214740
cctaaccaca gcgtttgtgc cttcagccgg ggaggcagcc cttgggacca gcatccctag 214800
ggacaggagg cagcgggaat catctctgta tctcgggttc tgcccagggg atgggcagac 214860
tctgccatct cttgagtgtt cgtttggaga agcctgagat gtggccctctg ctgccttctc 214920
actagttgca gtctatgtaa ataaggtcaa taaattcttt ggaagagcca cggagctgag 214980
tgaggctgtg ttgtgttttg ctttgccctag gctgggctca ggcagctctg cctcagcctc 215040
ccaaggagct ggggaactgg tatatgtcac tgtatatgtc actgtgcctg gcttatggct 215100
tggtttggct ttttttcaga tgggtctcaag tgccctcaggt tggccttgat cttgggatga 215160
ccttcctgct tgaaacagag tagtgggctt ataggcatga cccaccaggt ccaattttta 215220
ttttttaaag gcattgattt ttatacgtgt atggttgttt tgcccacttg tacatatgca 215280
caccatactt gtgtctggtc cctgcggagg tcagaagagg gcatcgggat cacctggaac 215340
cgaagttaat gaatggttat gagccacatc tcgatgtgta agattgaacc tggatccttt 215400
gcaagagcag ccagtgttct taccactga gccatctcta agccccacac ccagcttctt 215460
ttgatacaag gtctggtagc tcaaacttga tatgcagccg aggaggttga cctggatttc 215520
cctacctaacc ctcttctctc taccttccaa gtgctgatat tatacatagg catggatagt 215580
catgcccacc agtttgcctt gatggcacca gagtcaggaa agtccaaacc tggtagttgc 215640
aaacacagca agagggtaga ggcagccatt gtccctctggc tgccttggat acagagcttc 215700
tgggttgggt ggccttgggt cagttttccg aatgggtcac ccttggggaa agggaacact 215760
gctgaagagg tgggaccctg ggagggccgg cctccagctg ggtctctcca gccctgcct 215820
tggaacctag gctggaggga gccaaccagg atcctggact tgctacagtt aggtgaacag 215880
gctcctgcag cctccccttc ccttgggtag ctgtgggtgt ggtggtggtg gtggtggtg 215940
tggtggtggt ggtggtggtg gtgggggggg gggngnngnt

```

<210> 17
 <211> 473
 <212> PRT
 <213> Mus sp.

<400> 17
 Met Lys Arg Ala Ser Ser Gly Gly Ser Arg Leu Leu Ala Trp Val Leu
 1 5 10 15
 Trp Leu Gln Ala Trp Arg Val Ala Thr Pro Cys Pro Gly Ala Cys Val
 20 25 30
 Cys Tyr Asn Glu Pro Lys Val Thr Thr Ser Cys Pro Gln Gln Gly Leu
 35 40 45
 Gln Ala Val Pro Thr Gly Ile Pro Ala Ser Ser Gln Arg Ile Phe Leu
 50 55 60
 His Gly Asn Arg Ile Ser His Val Pro Ala Ala Ser Phe Gln Ser Cys
 65 70 75 80

Arg Asn Leu Thr Ile Leu Trp Leu His Ser Asn Ala Leu Ala Arg Ile
 85 90 95
 Asp Ala Ala Ala Phe Thr Gly Leu Thr Leu Leu Glu Gln Leu Asp Leu
 100 105 110
 Ser Asp Asn Ala Gln Leu His Val Val Asp Pro Thr Thr Phe His Gly
 115 120 125
 Leu Gly His Leu His Thr Leu His Leu Asp Arg Cys Gly Leu Arg Glu
 130 135 140
 Leu Gly Pro Gly Leu Phe Arg Gly Leu Ala Ala Leu Gln Tyr Leu Tyr
 145 150 155 160
 Leu Gln Asp Asn Asn Leu Gln Ala Leu Pro Asp Asn Thr Phe Arg Asp
 165 170 175
 Leu Gly Asn Leu Thr His Leu Phe Leu His Gly Asn Arg Ile Pro Ser
 180 185 190
 Val Pro Glu His Ala Phe Arg Gly Leu His Ser Leu Asp Arg Leu Leu
 195 200 205
 Leu His Gln Asn His Val Ala Arg Val His Pro His Ala Phe Arg Asp
 210 215 220
 Leu Gly Arg Leu Met Thr Leu Tyr Leu Phe Ala Asn Asn Leu Ser Met
 225 230 235 240
 Leu Pro Ala Glu Val Leu Met Pro Leu Arg Ser Leu Gln Tyr Leu Arg
 245 250 255
 Leu Asn Asp Asn Pro Trp Val Cys Asp Cys Arg Ala Arg Pro Leu Trp
 260 265 270
 Ala Trp Leu Gln Lys Phe Arg Gly Ser Ser Ser Glu Val Pro Cys Asn
 275 280 285
 Leu Pro Gln Arg Leu Ala Asp Arg Asp Leu Lys Arg Leu Ala Ala Ser
 290 295 300
 Asp Leu Glu Gly Cys Ala Val Ala Ser Gly Pro Phe Arg Pro Ile Gln
 305 310 315 320
 Thr Ser Gln Leu Thr Asp Glu Glu Leu Leu Ser Leu Pro Lys Cys Cys
 325 330 335
 Gln Pro Asp Ala Ala Asp Lys Ala Ser Val Leu Glu Pro Gly Arg Pro
 340 345 350
 Ala Ser Ala Gly Asn Ala Leu Lys Gly Arg Val Pro Pro Gly Asp Thr
 355 360 365
 Pro Pro Gly Asn Gly Ser Gly Pro Arg His Ile Asn Asp Ser Pro Phe
 370 375 380
 Gly Thr Leu Pro Ser Ser Ala Glu Pro Pro Leu Thr Ala Leu Arg Pro
 385 390 395 400
 Gly Gly Ser Glu Pro Pro Gly Leu Pro Thr Thr Gly Pro Arg Arg Arg
 405 410 415
 Pro Gly Cys Ser Arg Lys Asn Arg Thr Arg Ser His Cys Arg Leu Gly

420				425				430							
Gln	Ala	Gly	Ser	Gly	Ala	Ser	Gly	Thr	Gly	Asp	Ala	Glu	Gly	Ser	Gly
435				440				445							
Ala	Leu	Pro	Ala	Leu	Ala	Cys	Ser	Leu	Ala	Pro	Leu	Gly	Leu	Ala	Leu
450				455				460							
Val	Leu	Trp	Thr	Val	Leu	Gly	Pro	Cys							
465				470											

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
11 April 2002 (11.04.2002)

PCT

(10) International Publication Number
WO 02/029059 A3(51) International Patent Classification⁷: C12N 15/12,
C07K 14/705, 16/28, C12N 15/62, G01N 33/68, A61K
38/1706437 (US). CATE, Richard, L. [US/US]; 40B Nichols
Road, Cohasset, MA 02025 (US). SAH, Dinah, W., Y.
[US/US]; 4 Longfellow Place, Apt. 2608, Boston, MA
02114 (US).

(21) International Application Number: PCT/US01/31488

(74) Agents: HALEY, James, F. et al.; Fish & Neave, 1251
Avenue of the Americas, New York, NY 10020 (US).

(22) International Filing Date: 6 October 2001 (06.10.2001)

(25) Filing Language: English

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU,
ZA, ZW.

(26) Publication Language: English

(30) Priority Data:
60/238,361 6 October 2000 (06.10.2000) US(71) Applicants (for all designated States except US): YALE
UNIVERSITY [US/US]; 541 College Street, New Haven,
CT 06520 (US). BIOGEN, INC. [US/US]; 14 Cambridge
Center, Cambridge, MA 02142 (US).(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): STRITTMATTER,
Stephen, M. [US/US]; 96 Tulip Tree Drive, Guilford, CT

[Continued on next page]

(54) Title: NOGO RECEPTOR HOMOLOGS

Human NOGO-R1	1	-----MKRASA	GGSRLLAHWL	WLQAWQVAAP	CPGACVCYNE	50
Murine NOGO-R1		-----MKRASS	GGSRLLAHWL	WLQAWRVATP	CPGACVCYNE	
Murine NOGO-R3		MSWQSGTTVT	QSPVQAAQVS	GCCVELLLLL	LAGEFLPGGG	CPRDCVCYPA
Human NOGO-R3		-----	-----	-----	-----	
Human NOGO-R2		-----MLPGLRRLQ	APASACLLM	LLALPLAAPS	CPMLCTCYSS	
Consensus		-----	-----	-----	CP--C-CY--	

Human NOGO-R1	51	LRR NT		LRR 1	100	
Murine NOGO-R1		PKVTTSCPQQ	GLQAVPVGIP	AASQRIFLHG	NRISHVPAAS	FRACNLTIL
Murine NOGO-R3		PKVTTSCPQQ	GLQAVTPGIP	ASSQRIFLHG	NRISHVPAAS	FQSCNLTIL
Human NOGO-R3		P.MTVSCQAH	NFAAIPGEGIP	EDSERIFLQN	NRITFLQGGH	FSP..AMVTL
Human NOGO-R2		-----EGIP	-----	VDSERVFLQN	NRIGLLQPGH	FSP..AMVTL
Consensus		P..PTVSCQAN	NFSSVPLSLP	PSTQRLFLQN	NLIRTLRPGT	FGS..NLLTL
		P--T-SC--	-----P--P	-----R-FL--	N-I-----	F-----L

Human NOGO-R1	101	LRR 2		LRR 3	150	
Murine NOGO-R1		WLHSNVLARI	DAAAFGLAL	LEQLDLSDNA	QLRSVDPATF	HGLGRHLTLH
Murine NOGO-R3		WLHSNALARI	DAAAFGLTL	LEQLDLSDNA	QLHVVDPTTF	HGLGLHLTLH
Human NOGO-R3		WIYSNNITFI	APNTFEGFVH	LEELDLDGDN	QLRTAPETF	QGLVXLHLY
Human NOGO-R2		WIYSNNITYI	HPSTFEGFVH	LEELDLDGDN	QLRTAPETF	QGLVXLHLY
Consensus		W--SN----I	----F----	LE-LDL-DN-	-L----P-TF	-GL--L--L-

Human NOGO-R1	151	LRR 4		LRR 5	200	
Murine NOGO-R1		LDRCGLQELG	PGLFRGLAAL	QYLYLQDNAL	QALPDOTFRD	LGNLTHLFLH
Murine NOGO-R3		LDRCGLRELG	PGLFRGLAAL	QYLYLQDNAL	QALPDNTFRD	LGNLTHLFLH
Human NOGO-R3		LYKCGLSALP	AGIFGGLHSL	QYLYLQDNHI	EYLQDDIFVD	LVNLSHLFLH
Human NOGO-R2		LYKCGLSALP	AGVFGGLHSL	QYLYLQDNHI	EYLQDDIFVD	LVNLSHLFLH
Consensus		LYRCQLSSLP	GNIFRGLVSL	QYLYLQENSL	LHLQDDLPAD	LANLSHLFLH
		L--C-L--L-	---F-GL--L	QYLYLQ-N--	--L-D--F-D	L--NL-HLFLH

(57) Abstract: The invention relates generally to genes that encode proteins that inhibit axonal growth. The invention relates specifically to genes encoding NgR protein homologs in humans and mice. The invention also includes compositions and methods for modulating the expression and activity of Nogo and the NgR proteins. Specifically, the invention includes peptides, proteins and antibodies that block Nogo-mediated inhibition of axonal extension. The compositions and methods of the invention are useful in the treatment of cranial or cerebral trauma, spinal cord injury, stroke or a demyelinating disease.

WO 02/029059 A3



Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(88) Date of publication of the international search report:

23 January 2003

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 01/31488

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C07K14/705 C07K16/28 C12N15/62 G01N33/68 A61K38/17		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K G01N A61K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, BIOSIS, EMBL, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL NUCLEIC ACIDS 'Online! 21 January 2000 (2000-01-21) "M. musculus chromosome 11 clone RP23-113D4" retrieved from EMBL Database accession no. AC021768 XP002217414 compare translation product of complementary strand of nt 153209-154529 with aa 21-461 of seq. 4 and aa 1-392 of seq. 14 --- -/--	1,2,5-30
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *&* document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
18 October 2002		08/11/2002
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer
		Galli, I

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/31488

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL NUCLEIC ACIDS 'Online! 15 November 1999 (1999-11-15) "Homo sapiens clone RP11-10E2" retrieved from EMBL Database accession no. AC013606 XP002217419 compare aa 168-420 of seq. 2 with the translation product of nt 138578-139331.</p> <p>---</p>	1-4,6-30
X	<p>DATABASE EMBL PROTEINS 'Online! 1 May 2000 (2000-05-01) CORBY N.: "New leucine-rich protein" retrieved from EMBL Database accession no. Q9UGS3 XP002217415 compare aa 352-399 with seq. 11,19</p> <p>---</p>	1,11,15
A	<p>DATABASE EMBL PROTEINS 'Online! 1 October 2000 (2000-10-01) OSADA N. ET AL.: "Unnamed protein product" retrieved from EMBL Database accession no. Q9N0E3 XP002217416 compare with sequences 2,4,5,11,17,18,19</p> <p>---</p>	1-30
A	<p>WO 00 31235 A (CHEN MAIO S ;SCHWAB MARTIN E (CH)) 2 June 2000 (2000-06-02) the whole document</p> <p>---</p>	1-30
P,A	<p>WO 01 51520 A (UNIV YALE ;STRITTMATTER STEPHEN M (US)) 19 July 2001 (2001-07-19) NgR1: compare seq. 1,3 with seq. 5,17 of present application, respectively</p> <p>-----</p>	1-30

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims 1,2,6,7,9-13,15,16,18-30 relate to an extremely large number of possible compounds. It is not possible to determine all the possibilities, as any given X can be any amino acid or even a gap. In fact, the claims contain so many options, variables, possible permutations and provisos that a lack of clarity (and/or conciseness) within the meaning of Article 6 PCT arises to such an extent as to render a meaningful search of the claims impossible. Consequently, the search has been carried out for those parts of the application which do appear to be clear (and/or concise), namely seq. ID 2, 4 and 14.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/31488

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 27, 29 and, as far as methods in vivo are concerned, claims 26 and 28 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/31488

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 0031235	A	02-06-2000	AU 1469200 A	13-06-2000
			CN 1354755 T	19-06-2002
			CZ 20011608 A3	17-10-2001
			EP 1124846 A2	22-08-2001
			NO 20012223 A	02-07-2001
			SK 6222001 A3	03-12-2001
			WO 0031235 A2	02-06-2000
<hr/>				
WO 0151520	A	19-07-2001	AU 2940101 A	24-07-2001
			EP 1248803 A2	16-10-2002
			WO 0151520 A2	19-07-2001
			US 2002012965 A1	31-01-2002
			US 2002077295 A1	20-06-2002
<hr/>				